

## Importance and Interaction of Feedback Variables:

### A Model for Effective, Dynamic Feedback

ANNE COLLINS MCLAUGHLIN

WENDY A. ROGERS

ARTHUR D. FISK

Technical Report HFA-TR-0603

Atlanta, GA: Georgia Institute of Technology

School of Psychology - Human Factors and Aging Laboratory

February 2006



**Requests for more information may be sent to Wendy A. Rogers, School of Psychology,  
Georgia Institute of Technology, Atlanta, GA 30332-0170 ; e-mail: [wr43@mail.gatech.edu](mailto:wr43@mail.gatech.edu)**

## Table of Contents

<i>Executive Summary</i>	<b>4</b>
<i>Chapter 1 – Definitions and Constraints</i>	<b>7</b>
<i>Feedback Defined</i>	<b>7</b>
<i>Learning Defined</i>	<b>7</b>
<i>Similarities between Task Domains</i>	<b>8</b>
<i>Performance versus Learning</i>	<b>9</b>
<i>Common Feedback Terminology</i>	<b>10</b>
<i>Commonly Studied Feedback Variables</i>	<b>13</b>
<i>Chapter 2 – Models of Feedback</i>	<b>15</b>
<i>Response Certitude</i>	<b>15</b>
<i>Interference Perseveration</i>	<b>18</b>
<i>The Delay-Retention Effect (DRE)</i>	18
<i>The Interference Perseveration Hypothesis</i>	19
<i>Guidance Hypothesis</i>	<b>23</b>
<i>Self-regulated Learning</i>	<b>24</b>
<i>Interactive Cognitive Complexity Learning Model</i>	<b>26</b>
<i>Summary of Models of Feedback</i>	<b>27</b>
<i>Chapter 3 – Feedback Variables</i>	<b>31</b>
<i>Content</i>	<b>31</b>
<i>Precision</i>	<b>35</b>
<i>Timing</i>	<b>36</b>
<i>Trials-delay</i>	<b>38</b>
<i>Frequency</i>	<b>39</b>
<i>Absolute</i>	39
<i>Relative</i>	40
<i>Interpolated Activities</i>	<b>41</b>
<i>Chapter 4 – Learner Characteristics</i>	<b>43</b>
<i>Immutable Learner Characteristics</i>	<b>43</b>
<i>Cognitive Aptitudes</i>	44
<i>Working Memory Capacity</i>	44
<i>Mutable Characteristics</i>	<b>45</b>
<i>Experience</i>	45
<i>Arousal</i>	46
<i>Motivation</i>	48

<b>Chapter 5 – Task Demands</b>	<b>50</b>
<i>Complexity</i>	<b>50</b>
<i>Practice Schedules</i>	<b>51</b>
<b>Chapter 6 – Known Interactions</b>	<b>52</b>
<i>Aptitude x Treatment Interactions</i>	<b>52</b>
<i>Timing x Error Type Interactions</i>	<b>55</b>
<i>Ability x Motivation Interactions</i>	<b>56</b>
<i>Content x Error Type Interactions</i>	<b>57</b>
<i>Summary of Known Interactions</i>	<b>58</b>
<b>Chapter 7 – Principles of Feedback Design</b>	<b>59</b>
<b>Chapter 8 – A Conceptual Model of Effective Feedback</b>	<b>60</b>
<b>Chapter 9 – Conclusion</b>	<b>62</b>
<b>References Cited</b>	<b>63</b>
<b>Author Note</b>	<b>76</b>

## **Executive Summary**

This project combines the issue of understanding the role of the human operator (in collaboration with automated systems) with the interface technology issue of translating data into feedback. Automated systems can collect potentially unlimited amounts of data. A critical challenge is ensuring that those data are presented properly to the human when needed for appropriate management of, and learning about, the automated system. Variables such as the content and the timing of this feedback information must be understood. The goal of this report is to provide such understanding and to provide insight about when feedback supports learning, when it interferes with learning, and how optimality of feedback differs across tasks, contexts, and users.

Feedback is information. Feedback can be used to correct errors, to monitor a system's state, and to support learning. Throughout the past century, researchers and designers have tried to discover how humans learn and how a system can support that learning and help correct mistakes. Systems may contain inherent feedback (such as the sound of an engine), but most require augmented feedback to be provided by the system designer. Error messages are an example of designer-provided feedback. Feedback may concern the operator's performance, the state of the system, or monitoring information about the system. Feedback may contain different amounts and types of information and occur throughout the use of a system. The purpose of this review was to identify recommendations and best practices for the design of feedback.

Based on a systematic, integrative analysis of the research literature, we developed a conceptual model of the critical factors that affect feedback efficacy. The three main components of the model are:

**(1) Learner characteristics**

- a. **User abilities** – characteristics of a person such as intelligence, working memory capacity, and spatial ability. Feedback should be calibrated to a learner's abilities.
- b. **Current state of a user** – fatigue, arousal, and motivation levels; designers need to understand how the success of feedback may *interact* with and *change* these states.

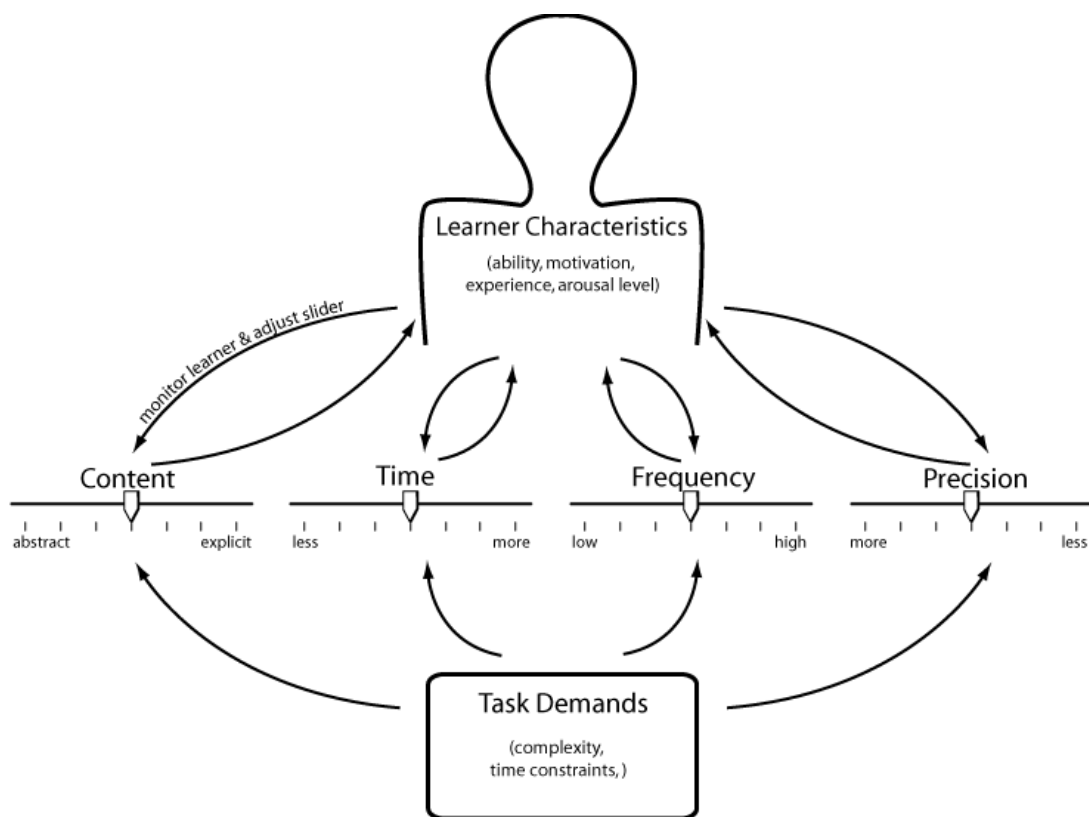
**(2) Task demands**

- a. **Simple** – tasks that have few components or impose minimal working memory demands on the learner. Such tasks may leave resources free for comprehending feedback.
- b. **Complex** – tasks that impose more mental workload on the learner, may lead to feedback being missed or ignored.

**(3) Feedback characteristics**

- a. **Content** – feedback must be matched to the amount of information a learner can handle for a particular task to ensure it is not followed mindlessly, misunderstood, or ignored.
- b. **Timing** – feedback must be calibrated to ensure the information is relevant and that the user is able to process the information at that point in time.
- c. **Frequency** – whether feedback is presented every time an event occurs or on a different schedule.
- d. **Precision** – the specificity of the information provided by the feedback.

The model below illustrates that feedback effectiveness is influenced both by the characteristics of the learner and by the demands of the task to be learned. Our analysis thus far suggests that learning through feedback depends on multiple learner characteristics, including ability and prior experience. It is not merely a question of what content to present when. The degree to which the content is informative to support system use and learning will depend on to whom it will be presented, as well as on what type of task is being learned or performed. The details of the model are presented in Chapter 8.



To summarize, this report supplies an understanding of feedback as studied in psychological research, a review and critique of past models of feedback, and a specification of the commonly studied feedback variables and their interactions. In addition, our review provides insight into the overall picture of how feedback design affects learning, as summarized by the model above.

## Chapter 1 – Definitions and Constraints

### *Feedback Defined*

What is feedback? There is not a simple straightforward answer to this question. The concept of feedback spans numerous areas, from electrical engineering to management to psychology. Indeed, even in psychology the concept of feedback is defined in different ways. For example, in the behaviorist tradition, feedback is reinforcement. Giving feedback increases a particular behavior, either through positive or negative reinforcement. In management and organizational psychology, feedback generally means to let someone know “how they are doing.” A trainee performs a job and feedback from a superior is intended to improve and maintain performance. In cognitive and educational psychology, feedback is information. Feedback can be used to correct errors or to provide information for learning.

A constraining definition was necessary to focus our review on issues related to human-system interactions. The definition of feedback we use was **augmented, external information meant to promote human learning** (derived from Salmoni, Schmidt, & Walter, 1984).

Augmented means the feedback must contain more information than was present inherent to the task. For example, if the task were to press a button, feedback was not simply the button depressing, there must be *additional* information contained for it to be considered feedback. External meant that the information must come from outside the learner, that is, provided by another person or system. This definition had a cognitive slant, in that *learning* was specified rather than behavior and *information* rather than reinforcement.

### *Learning Defined*

Many of the mixed findings concerning effective feedback may be traced back to how

learning was defined in each study. Some considered learning to be the ability to perform a task (with or without feedback in place; Bilodeau & Bilodeau, 1958), whereas others more strictly defined learning as performance on a retention or transfer test with no feedback present (Salmoni et al., 1984). We define learning as **a relatively permanent change in behavior brought about by experience** (adapted from Salmoni et al., 1984). Experience may include direct engagement, observation, or feedback. However, feedback must be removed to test learning.

### *Similarities between Task Domains*

Research on skill acquisition was traditionally divided into the categories of motor skills, psychomotor skills (such as line drawing), simple cognitive skills (such as associative learning), and complex cognitive skills (such as problem solving). We reviewed research in all three areas to assess generalities along the spectrum from motor to cognitive tasks. Previous research has generally focused primarily on one area or another, keeping overall patterns hidden.

In his 1987 review of learning, retention, and transfer of human motor skills, Adams defined skill as having three characteristics: it contained complex behavior, was learned, and had components that were cognitive, perceptual, and motor. A skill might be heavily weighted in one category, but contained all three components. A recent review suggested that the acquisition of psychomotor and intellectual skills followed similar patterns (Rosenbaum, Carlson, & Gilmore, 2000). Research on simple and complex cognitive tasks often relies on work from the psychomotor domain to formulate principles of learning and feedback (Goldstone, 1998).

Perceptual-motor skill learning shared the power law function with numerous cognitive skills and contributed to the goal of a general theory of learning (Newell, 1991). Other similarities between domains include warm-up effects, contextual interference effects, and



randomized versus blocked practice effects (Carlson & Yaure, 1990; Proctor & Dutta, 1995; Schneider, Healy, & Bourne, 2002; Simon & Bjork, 2001).

The similarities between task domains do not mean there are not meaningful differences across domains. For example, definitions of feedback timing indicated differences in the empirical methods between domains. In psychomotor skill acquisition, responses were often a repetition toward a particular goal action. Therefore, the goal of the next response was the same as that of the previous response, though presumably closer to the goal state. In the cognitive or verbal learning domains, “responses” were often answers to questions and each required a different “correct” answer. Therefore, although the goal for both was to produce the correct response, the correct response changed with every trial in cognitive skill acquisition. This methodological difference should be kept in mind when considering feedback design.

### ***Performance versus Learning***

The early literature on feedback had notoriously mixed findings (Salmoni et al., 1984). In most of the early studies, researchers did not adequately differentiate between learning and performance. Performance was measured with feedback present but referred to as learning. However, performance with feedback present was not a demonstration of learning when learning was defined as a *relatively permanent change* (Salmoni et al., 1984). Improvements in performance with feedback present was learning to perform *with* the feedback. It was not necessarily learning the task or skill. Most early research did not include either a retention or transfer test (with feedback removed); thus the mix of findings was perhaps not surprising. More recently, researchers have been likely to include retention or transfer as a measure of learning.

### *Common Feedback Terminology*

Feedback terms have been used fairly consistently over the last 100 years in reference to feedback research. These terms are presented in Table 1 (alternate names are listed in the rightmost column).

Listed first in Table 1 is internal feedback which differs from other feedback in that it is not presented to the learner but comes from within. For example a diver knowing that a movement was incorrect or a tractor operator sensing a change of machine vibration that might indicate a system malfunction. Developing internal feedback is crucial to demonstrating learning at retention and transfer, once external feedback is removed. The ability to self-correct internally without the presentation of augmented, external feedback is a demonstration of learning.

Second in the table is the most studied form of feedback as information – knowledge of results (KR). KR was previously referred to as reinforcement, but now defined as information (see Mory, 2004, for a review). KR only provides information about the correctness of a response: right or wrong. Part of its appeal was that it could be isolated from other forms of feedback and developed for almost any task. One way KR was applied in the instructional literature was “Answer Until Correct” (AUC). This meant to provide KR on a trial after each response, not progressing to a different trial until the learner made the correct response.

The third type of feedback in Table 1 is knowledge of correct response (KCR) which provides more information than KR as it contains the correct response. Instead of saying choice A on a test was incorrect, KCR feedback would also state that choice C was correct. KCR was more informative than KR and provided more structure, as the learner was informed of the correct answer. A sub-category of KCR is amount of information (I) provided along with the correct response. Often called elaborative feedback, this information ranged across studies in

Table 1

*Definitions of Feedback Terminology*

<b>Term</b>	<b>Definition</b>	<b>Example</b>	<b>Other names</b>
Internal feedback	Originates with the learner, not guaranteed to be correct.	Diver: "That first twist was off."	
Knowledge of results (KR)	Contains only accuracy information and is provided outside the response-produced feedback of the task. Answer until correct (AUC) - This is a form of KR studied in the instructional literature. Learners respond, receive KR (not KCR), and must respond again until correct.	"right / wrong", "correct / incorrect"	Minimal feedback, augmented feedback, outcome feedback
Knowledge of correct response (KCR)	Contains the correct response, implicitly or explicitly contains KR as well.  Information (+I) - May be added to KR or KCR. Information indicates elaboration beyond KR or KCR. The type of elaboration must be stated.	"the correct answer was A" , "The line should be X long." "The answer was A, because 'wren' is a type of bird." "Answer C is correct because deer are mammals. They are related to elk and zebras."	Corrective feedback  Informative feedback, elaborative feedback, topic contingent, response contingent
Performance feedback	Contains two subtypes sometimes defined differently and sometimes used interchangeably. Kinetic - feedback on the forces, spatial, and temporal properties used in motion, differs from outcome KR as it is not only the eventual KR of the goal, but is on the forces important to achieve the goal. Kinematic - refers to specifics of the movement to produce the outcome.	Dive instructor points out diver did not jump high enough from the board, which contributed to failure in dive as a whole. Dive instructor points out twist in a dive that was not completed.	Knowledge of performance, kinetic feedback, kinematic feedback
Summary feedback	Feedback occurring after response is completed. May be KR or summative feedback over a number of trials.	A graph of responses over 100 trials	Terminal feedback, trials-delayed feedback

level of detail, structure, and precision. However, because it was not consistently defined it is difficult to draw conclusions about the boundaries of elaborative effectiveness.

The fourth listing in the table, performance feedback, is defined as information given on the process that *led* to a result. Simply knowing whether that result was correct or not would be KR, but an example of performance feedback would be a dive instructor who pointed out to a diver that she did not jump high enough to complete the dive correctly. Or a commercial mower operator being shown why a cut was not adequate. Performance feedback often produced greater learning than KR (e.g., Buff & Campbell, 2002; Swinnen, Lee, Vershueren, Serrien, & Bogaerds, 1997). However, this may be due to *additional information present* in performance feedback rather than performance feedback as a unique category.

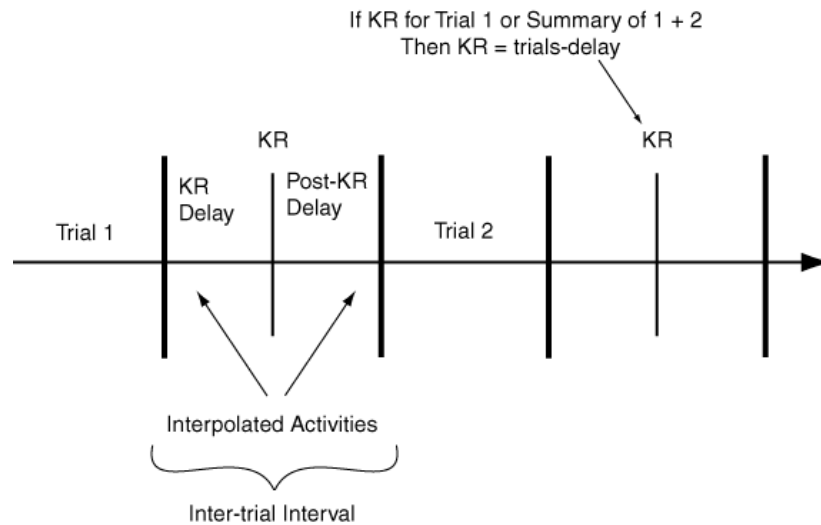
The next table entry, summary feedback has mostly been investigated in psychomotor learning. This type of feedback provides a summary of performance over a number of trials. For example, if the task were to learn to apply a certain amount of pressure to a transducer, summary feedback might be a graph of the last ten trial responses compared to an indicator of correctness. Summary feedback usually contains KR, but might also contain KCR or other feedback types. For example, summary feedback in a study of bimanual coordination contained a replay of the trial; this was performance feedback *and* summary feedback (Wishart et al., 2002). Because there were a number of ways to present summary feedback, it was difficult to compare studies.

In summary, there are different types of feedback and we will use the terminology presented in Table 1 for the remainder of the report. As will become evident, the literature does not provide clear guidance about which type of feedback is best. Part of the problem lies in methodological differences across studies.

### *Commonly Studied Feedback Variables*

The feedback variables that have been extensively examined include: frequency of feedback presentation, delay of feedback, amount of time between feedback and the next required response, inter-trial interval, activities taking place within that interval, amount of trials before feedback presentation, and precision of the feedback.

Frequency of feedback is how often feedback is presented during learning. Absolute frequency refers to the total number of feedback instances, whereas relative feedback reflects the percent of trials on which feedback is given. Delay of feedback is the amount of time between learner response and feedback presentation and post-feedback delay is the amount of time between feedback presentation and initiation of the next trial (see Figure 1).



*Figure 1.* Time frame of commonly studied feedback variables within trials of learning.

The inter-trial interval is the expanse of time between trials and includes feedback delay and post-feedback delay. Any activities other than feedback during that time are known as interpolated activities. If there are multiple trials before feedback is presented, it is said to be trials-delayed. Precision refers to how precisely the feedback corrects the response. Table 2 provides an overview of the effects of these variables on motor performance and learning.

Table 2

*Summary of Various Effects of Feedback on Performance and Learning (Adapted from Salmoni, Schmidt, & Walter, 1984, p. 378)*

<b>Increases in variables</b>	<b>Effects on motor performance</b>	<b>Effects on motor learning</b>
Absolute frequency of feedback	Enhanced	Enhanced
Relative frequency of feedback	Enhanced	Degraded
Slight delay of feedback	None or slightly degraded	None or degraded
Post-feedback delay	Enhanced slightly	Enhanced, if KR delay constant
Inter-trial interval	Mixed, unclear	Enhanced, if KR delay constant
Interpolated activities in feedback delay	Degraded, if demanding	Degraded
Interpolated activities in post-feedback delay	Degraded, if demanding	None
Trials-delay	Degraded	Enhanced
Precision	Enhanced	Enhanced

## Chapter 2 – Models of Feedback

Cognitive models of feedback are not as numerous as one might imagine. This seems due mostly to the fairly recent switch from modeling feedback as a reinforcer to modeling it as information processed by the learner. Several common models are the model of response certitude (Kulhavy & Stock, 1989), the interference-perseveration hypothesis (Kulhavy & Anderson, 1972), and the guidance hypothesis (Salmoni et al, 1984). Others have more recently appeared, such the model of self-regulated learning (Butler & Winne, 1995) and the interactive cognitive complexity model (Tennyson & Breuer, 1997).

### *Response Certitude*

The model of response certitude included many feedback variables: learner confidence, feedback content, how to correct errors, and the timing of the feedback (Mory, 2004). In the certitude model of feedback, three cycles defined an instructional episode (Kulhavy & Stock, 1989): (1) a stimulus requires a response; (2) the learner receives feedback; (3) the learner uses that feedback to respond again to the stimulus from the first cycle (this assumes closed-loop learning). The application cycle of response certitude is in Table 3.

The first step is to provide a practice question to assess the current knowledge state of the learner; the learner selects a response with an amount of certitude. Then, the learner receives feedback and processes the discrepancy between the feedback and internal response certitude. If the feedback validates an initial high-certitude response, little processing is necessary. The correct answer is known. If the feedback corrects an initial high-certitude response, the high discrepancy predicts more in-depth processing of that feedback. If the learner was not confident of the response and received feedback of correctness, this indicates low discrepancy.

Table 3

*Model of Response Certitude (Adapted from Dempsey, Driscoll, & Swindell, 1993)*

	Input	Learner	Output
Cycle I	Practice question	1. compared question to current knowledge 2. evaluated response possibilities 3. selected response possibilities	Responded to question
Cycle II	Feedback	1. compared to question and response given 2. modified confidence of response based on perceived discrepancy	Studied feedback
Cycle III	Posttest question	1. compared again to knowledge base 2. selected a response with associated certitude	Responded to posttest



The discrepancy value was calculated as:  $f(v) * c = d$ , where  $f(v)$  was the verification component,  $c$  was certitude, and  $d$  was discrepancy between the two. Discrepancy ranged from -5 to +5 and predicted how much time spent studying the feedback (conceptualized as effort) the learner put into error correction. Effort would be a better term, if more difficult to measure, as some stimuli may be processed faster *and* more deeply than others ( Craik & Lockhart, 1972), but effort has been typically operationally defined as time spent studying the feedback. The relation of effort to discrepancy had been validated by data; learners spent more time studying feedback when discrepancy was high (Kulhavy & Stock, 1989; Matsui, Kakuyama, & Onglatco, 1987). Indeed, even when learners were told they were incorrect after giving a *correct* answer, similar results were found (Kulhavy & Stock, 1989). Learners spend more time studying feedback when told their answer was incorrect, regardless of actual accuracy.

The discrepancy between a learner's initial certainty of response correctness and feedback predicted the amount of feedback processing effort a learner put forth. This was similar to the closed-loop theory of learning in that correction and eventual correct performance occurred through a comparison between desired and initial state. The important difference in the response-certitude model was that certitude drove the depth of processing during the comparison and eventual correction. According to the response certitude model, the amount of elaboration in feedback content should be driven by the discrepancy value. Low discrepancy, as when a learner gave a correct response which he or she knew was correct, required little elaboration. High discrepancy, when an incorrect response was paired with high certitude of correctness, required the most elaboration, and the learner put forth the most effort in processing this elaboration. This related to the idea that humans try to reduce discrepancy between their actual and desired actions (Kanfer & Ackerman, 1989). The processing of feedback, whether at shallow or in-depth levels,

affected the response to the item on a retention test. This likely interacted with learner ability, where high-ability learners might put forth effort and have the resources to do so. Low-ability learners and those learning difficult tasks, (the most likely to misjudge their own responses in the first place (Dunning, Johnson, Ehrlinger, & Druger, 2003; Metcalfe, 1986; Mory, 1994), would not have the resources to understand the elaborative feedback, even when the discrepancy value predicted they would put forth more effort.

### ***Interference Perseveration***

Asking a learner to estimate the certainty of a response before receiving feedback was a self-evaluative interpolated activity. An interpolated activity could take place either after a trial but before feedback, or after feedback but before the next trial. Investigation into the effects of interpolated activities earned them a place as one of the feedback variables listed in detail in the next section. For now, it is enough to know that resource demanding interpolated activities tended to reduce learning, unless the interpolated activity was self-evaluation.

### ***The Delay-Retention Effect (DRE)***

Seemingly against the findings for self-evaluative interpolated activities was the delay-retention effect (DRE). The possibility that delaying feedback was beneficial to learning had to be addressed after the early findings of Brackbill, Bravos, and Starr (1962). In this first study to find the delay retention effect, feedback was KCR and the immediate feedback was given as soon as a response was made. For the delayed feedback groups, the initial stimulus was left up for either 5 or 10 seconds until KCR. In essence, the delayed feedback group received presentation of the initial test item (stimulus) during the entire delay interval, whereas the

immediate group received the initial test item briefly followed by KCR. Despite confounded stimulus presentation length, this study was immediately hailed as a fascinating finding for the positive effects of delaying feedback.

Kulhavy and Anderson (1972) examined the Brackbill et al., study and hypothesized that the answer just given by the learner was still in mind when the feedback appeared with no delay. Conceptually, the slate on which the initial incorrect response was written was wiped clean by a delay, and allowed clear writing of the correct response, rather than one clouded by the prior response. Therefore, a previous incorrect answer interfered with the feedback and the feedback was not encoded. This was termed the *interference perseveration* hypothesis .

### ***The Interference Perseveration Hypothesis***

According to the interference perseveration hypothesis, given a delay, the error response made by the learner faded and when the correct response was presented, the learner strengthened association between the question and the correct response (Kulhavy & Anderson, 1972; Lavery, 1962; Sassenrath & Yonge, 1968; Swinnen et al., 1990). A meta-analysis of feedback timing studies found delayed feedback resulted in 0.44 standard deviations higher scores than immediate feedback (Kulik & Kulik, 1988). Thus an interpolated activity would serve the same purpose of fading the initial error response before receiving feedback.

However, there were substantive standardization issues with these studies and likely a social reason behind their conclusions as well. In 1967, philosophy of science was moving away from behaviorism toward cognitivism. One basic tenet of behaviorism was the need to closely pair stimulus and feedback in time. It was possible that any evidence against this paradigm was taken as confirmation that behaviorism could not explain the full functioning of the mind.

In other studies, “immediate” and “delayed” feedback had differing definitions. For example, in Kulhavy and Anderson (1972), immediate feedback occurred after the 35-item multiple-choice test was completed. Delayed feedback consisted of the test and answers presented a day later. In earlier literature, “immediate” indicated after trial feedback, not after test (Salmoni et al., 1984). The following studies specified exactly how delay was manipulated.

In a motor learning task, delaying feedback by several seconds (3.2 s) after each trial resulted in better retention than instantaneous feedback (210 ms after each trial; Swinnen et al., 1990). In this study, the pre-feedback interval contained no interpolated activities. Similarly, when the interval was increased to eight seconds for delayed feedback and a condition was introduced where one group estimated their own errors during those eight seconds, that group appeared to show better retention (Figure 2), but the difference was not significant.

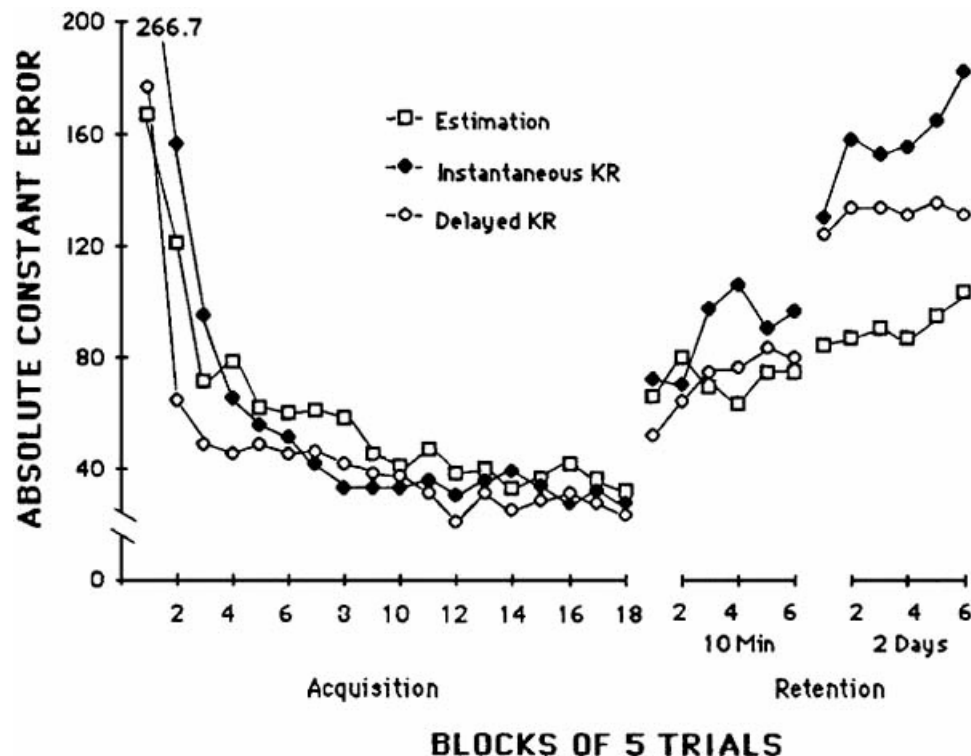


Figure 2. Data reprinted from Swinnen, et al., (1990). X-axis shows blocks of trials across time, then retention scores (with no feedback present) after 10 minutes and 2 days. Y-axis shows amount of absolute constant error on a motor learning task.

Sassenrath and Yonge (1968) tested the effects of 24-hours of delay. Those in the immediate feedback condition took a test of introductory psychology and then were immediately given the same test with the correct answers included. They then re-took the original test, with re-ordered questions. The delayed group received the test with the answers 24-hours later and then took the re-ordered test. Both groups performed similarly on the second test. To test retention, they again took the test five days after their first testing day. On this test, the delayed feedback group outperformed the immediate feedback group (Sassenrath & Yonge, 1968). This study supported the interference perseveration hypothesis.

A second finding by Kulhavy and Anderson (1972) was that learners who received delayed feedback studied the feedback for a longer time. The authors proposed the delay resulted in increased attention to the feedback, as there was no longer interference from the original trial. The finding of increased study time occurred in many studies that found a delay retention effect and supported the notion that the delay time was used for self-assessment.

Some evidence against the interference perseveration hypothesis was that other studies showed improvements in learner performance on a retention test, yet those learners could recall their initial wrong answers when queried (Peeck & Tillema, 1979), which disproved that perseveration of errors and learning cannot co-exist.

It was likely that the delayed feedback recipients in most DRE studies *re-familiarized* themselves with the subject matter and then the answer, thus the increase in study time. This could explain the DRE. Learners with delayed feedback got more exposure to the material. They received one exposure trial when answering the question and a second when re-familiarizing themselves with the question and answer. The immediate feedback group did not need to re-familiarize themselves with the question because they just saw it. Therefore, they only looked at

it once, looked at the feedback once, and had effectively fewer trials. There was support for this hypothesis in the motor learning literature. Lee and Magill (1983) found that systematically varied practice schedules (not random) for a motor task produced similar learning to random schedules (both more effective than blocked practice). More learning occurred because as the task changed between trials, (though the next task was known to the learner), re-familiarization had to occur. In similar research, practice schedules with demanding interpolated activities showed similar results (Lee & Magill). This same re-familiarization could occur after a delay of feedback and would constitute a double-stimulus exposure.

This confound of the double-stimulus was not addressed in the literature. However, studies controlling for the double-stimulus eliminated the DRE. For example, in Gaynor (1981), the delayed feedback format was “The answer for frame \_\_\_\_ was \_\_\_\_.” Not surprisingly, no differences were found for learners of similar skill and immediate feedback was superior for low skill learners. Other studies without a double stimulus exposure had “delays” of less than a minute, with no intervening events (see Kulik & Kulik, 1988, for review). Again, this pointed to problems classifying studies of feedback timing.

Another typical confound in studies finding the DRE was that the delay brought the viewing of the answers (feedback) closer in time to the post-test. For example, Swindell and Walls (1993) provided an initial test, delayed feedback by one day, and re-tested on the third day. Those receiving delayed feedback saw the questions and answers 66% closer in time to the retention test and did perform better than those who received feedback immediately after the initial test. Kulhavy and Anderson (1972) provided delayed feedback one day after the initial test, meaning those who received immediate feedback retained the information for seven days while the delayed group retained it for six. Sturges (1978) used a similar method and set the

retention test to anytime between seven and twenty-one days after the initial test. In each of these studies, the percent change in the immediate versus delayed group decreased. However, none of the studies made any mention that this might in some way account for the DRE.

One last problem with typical studies showing the DRE was cohesiveness of tested material. It was unknown whether knowing the answer to one question affected answers to other questions when test items were related (typical in a test of introductory psychology, etc.) When “immediate” feedback means after-item (in the inter-trial interval) rather than after-test, receiving feedback on one item might affect correctness on subsequent items. For example, if the answer to one question was known or revealed via immediate feedback, if that same answer were available on a subsequent question, it may be eliminated as a choice. This has not been investigated as potential noise in the feedback timing data.

### ***Guidance Hypothesis***

The general role of feedback is to guide and motivate learning (Salmoni et al., 1984). Salmoni et al. developed this idea after reviewing 250 studies on the effects of feedback in learning motor skills. Their re-categorization of these studies served as the basis for the guidance hypothesis. Learners continued to practice, “try harder,” and “be more interested in the task” when *able* to perform the task thanks to the guiding function of feedback (p. 379). On trials with no feedback, learners paid attention to internal sensory information to develop internal guidance for eventual demonstration of learning. Consequently, effects of feedback are similar to the effects of guided training. Keeping someone on task through strictly guided training resulted in a strong boost to performance during training with a severe drop-off when guidance was removed (Anderson, 1996; Salmoni et al.).

The guidance hypothesis pervaded the feedback literature after 1984 and was often misinterpreted to mean anything that makes feedback harder to understand will improve learning (Wulf, Shea, & Matschiner, 1998). The original article did make reference to the similarities of feedback and practice effects, where more “difficult” random practice resulted in more learning than blocked practice (Shea & Morgan, 1979), but there are many ways to make feedback harder to understand. Giving feedback in a foreign language and forcing the learner to translate it via a dictionary, for example, would probably not result in better learning of the task, yet it makes feedback harder to understand. Only feedback that guides the development of internal feedback will improve learning.

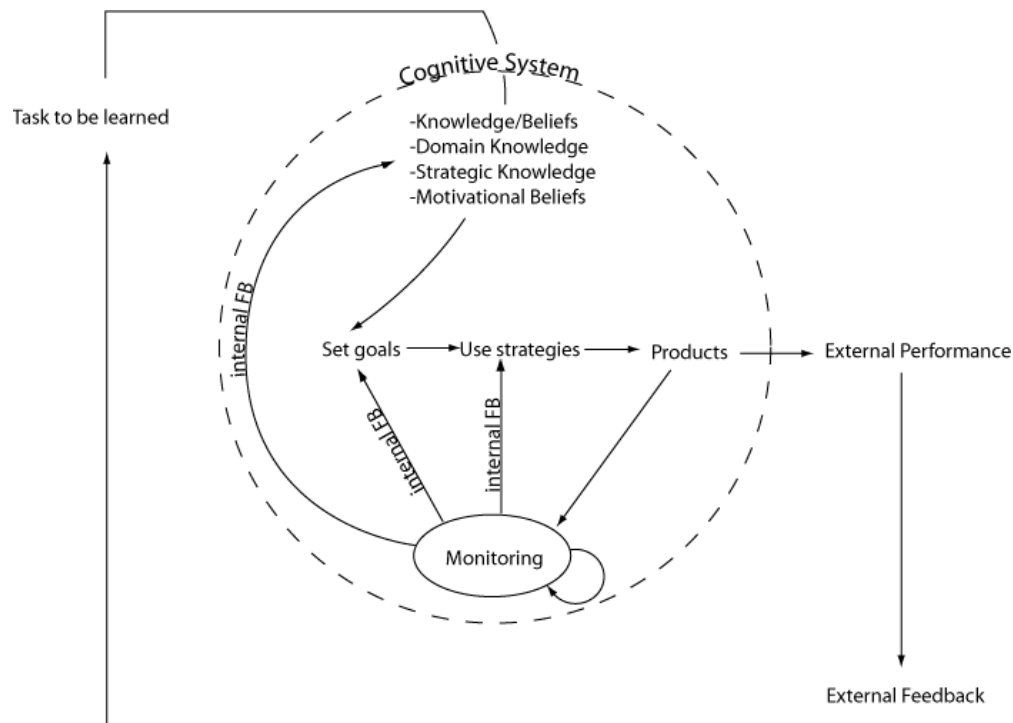
In essence, the guidance hypothesis provided the information as to *why* feedback variables showed different effects under different conditions. The next model discussed, self-regulated learning, contributed the idea of *what* occurs in a cognitive system guided by feedback.

### ***Self-regulated Learning***

The model of self-regulated learning was developed to explain the activities of effective learners (Butler & Winne, 1995). It was a closed-loop model where learners attempted a task, took feedback into account, and adjusted their actions. As seen in Figure 3, the task to be learned was filtered through the knowledge and beliefs of the learner to set goals, apply strategies toward those goals, and produce products (both cognitive and behavioral.)

During this process, the learner monitored all of these components to change them if necessary, producing internal feedback. External feedback produced the same result, but was also filtered through the knowledge and beliefs of the learner to enter the cognitive system. As illustrated in Figure 3, the system was entirely dynamic, with each part of it able to change





*Figure 3.* Model of self-regulated learning, adapted from Butler and Winne (1995). Task to be learned passes through the knowledge and beliefs of the learner to set goals, apply strategies toward those goals, and produce cognition and behavior. External feedback produces the same result, but was also filtered through the knowledge and beliefs of the learner to enter the cognitive system. The system is entirely dynamic, with each part able to change during learning.

during the learning process. It was similar to the components of the closed-loop theory of motor learning. For example, as the external task was attempted, the learner used internal feedback as well as external feedback in altering knowledge, strategies, and goals to accomplish the task. The more change to this cognitive system, the more learning occurred for the task.

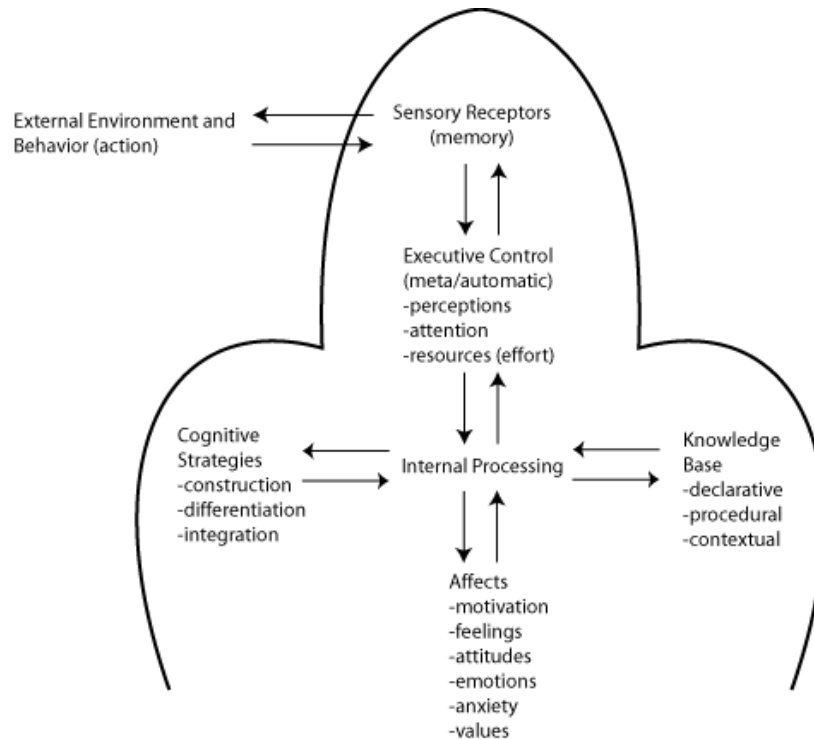
An issue with the self-regulated learning model was that it predicted a higher amount of elaborative, external feedback would best aid the learner in self-regulation. However, there was a large amount of evidence to the contrary; less external feedback (such as KR) often resulted in *more* learning and had been proven to be more effective than elaborated feedback (Birenbaum & Tatsuoska, 1987; Cope & Simmons, 1994; Goodman, Wood, & Hendrickx, 2004; Richards,

1989). Simply stated, more feedback or more elaborate feedback did not necessarily produce better learning and might actually inhibit learning under certain conditions.

### ***Interactive Cognitive Complexity Learning Model***

One trend in the philosophy of science was complexity theory, where mechanisms of the mind ran in parallel and adapted constantly to a changing environment. Tennyson and Breuer (1997) developed a complexity model (presented in Figure 4) of the learning system with specific feedback prescriptions based on the type of learning desired. The model was created to satisfy three requirements: account for linear and non-linear elements of learning, address the interaction of knowledge and strategy use in complex cognition, and include affective processes such as motivation and anxiety. The model relied on Gagne's (1970) varieties of learning outcomes which are presented in Table 4.

The model assumed several interactive systems including a sensory register (where environmental information enters the system), an executive control that controls perception, attention, and resources (i.e., working memory), a knowledge base (i.e., long term memory), an affective component for attitudes and motivation, and an internal processor. The internal processor interacted with most of the other systems to construct, integrate, and differentiate knowledge. This model incorporates many variables known to affect feedback efficacy such as task demands, internal and external feedback information, prior knowledge, learner ability, motivation, and arousal. Although the model does not provide clear guidelines of how those variables interact, its parallel, complex nature aided in our development of a model of feedback



*Figure 4.* The interactive cognitive complexity learning model, adapted from Tennyson and Breuer (1997). Information enters the cognitive system through sensory receptors; these receptors communicate with the executive control. Internal processing then interacts with cognitive strategies, knowledge base, and affect to inform the executive control and sensory receptors to produce behavior.

### ***Summary of Models of Feedback***

The purpose of reviewing current models of feedback was to understand how different researchers conceptualized feedback acting on a learner's cognitive system. Response certitude outlined how an interpolated activity could aid learning *because it focused the learner on developing internal feedback*. The interference perseveration hypothesis suggested how delay of feedback might serve to *focus the learner on developing internal feedback*. In the guidance hypothesis, feedback hardly affected learning at all, except that it encouraged and guided practice to develop the necessary *internal feedback*. In self-regulated learning, the most change to the *internal cognitive system* produced the most learning. In the interactive cognitive

complexity model, feedback interacted with the *internal processes* that eventually produced learned behavior. Despite differing assumptions, theoretical bases, and some differing prescriptions, these models had commonalties.

Table 4

*Gagne's varieties of learning outcomes (1970).*

Taxonomy of Learning Outcomes	Example	Critical Learning Conditions
<b>Verbal Information</b>	Stating previously learned materials such as facts, concepts, principles, and procedures, e.g., listing the seven major symptoms of cancer	<ol style="list-style-type: none"> <li>1. Draw attention to distinctive features by variations in print or speech.</li> <li>2. Present information so that it can be made into chunks.</li> <li>3. Provide a meaningful context for effective encoding of information.</li> <li>4. Provide cues for effective recall and generalization of information.</li> </ol>
<b>Intellectual Skills:</b> Discriminations, Concrete Concepts, Defined Concepts, Rules, Higher Order Rules	<p>Discriminations: Distinguishing objects, features, or symbols, e.g., hearing different pitches played on a musical instrument</p> <p>Concrete Concepts: Identifying classes of concrete objects, features, or events, e.g., picking out all the green M&amp;Ms from the candy jar</p> <p>Defined Concepts: classifying new examples of events or ideas by their definition, e.g., noting "she sells sea shells" as alliteration</p> <p>Rules: Applying a single relationship to solve a class of problems, e.g., calculating the earned run averages (ERA) of the Atlanta Braves</p> <p>Higher Order Rules: Applying a new combination of rules to solve a complex problem, e.g., generating a balanced budget for a state organization</p>	<ol style="list-style-type: none"> <li>1. Call attention to distinctive features.</li> <li>2. Stay within the limits of working memory.</li> <li>3. Stimulate the recall of previously learned component skills.</li> <li>4. Present verbal cues to the ordering or combination of component skills.</li> <li>5. Schedule occasions for practice and spaced review.</li> <li>6. Use a variety of contexts to promote transfer.</li> </ol>

Table 4 (cont.)

*Gagne's varieties of learning outcomes (1970).*

<b>Taxonomy of Learning Outcomes</b>	<b>Example</b>	<b>Critical Learning Conditions</b>
<b>Cognitive Strategies</b>	Employing personal ways to guide learning, thinking, acting, and feeling, e.g., devising a corporate plan to improve customer relations	<ol style="list-style-type: none"> <li>1. Describe or demonstrate the strategy.</li> <li>2. Provide a variety of occasions for practice using the strategy.</li> <li>3. Provide informative feedback as to the creativity or originality of the strategy or outcome.</li> </ol>
<b>Attitudes</b>	Choosing personal actions based on internal states of understanding and feeling, e.g., deciding to exercise daily as a part of preventive health care	<ol style="list-style-type: none"> <li>1. Establish an expectancy of success associated with the desired attitude.</li> <li>2. Assure student identification with an admired human model.</li> <li>3. Arrange for communication or demonstration of choice of personal action.</li> <li>4. Give feedback for successful performance; or allow observation of feedback in the human model.</li> </ol>
<b>Motor Skills</b>	Executing performances involving the use muscles, e.g., doing a triple somersault dive off the high board	<ol style="list-style-type: none"> <li>1. Present verbal or other guidance to cue the executive subroutine.</li> <li>2. Arrange repeated practice.</li> <li>3. Furnish immediate feedback as to the accuracy of performance.</li> <li>4. Encourage the use of mental practice.</li> </ol>

Note: Information from: Driscoll, M.(1991) *Psychology of Learning for Instruction*. Allyn and Bacon.

## Chapter 3 – Feedback Variables

### *Content*

Numerous subtypes of feedback content have been defined (Table 1). The types of feedback content may be thought of as a spectrum, where knowledge of results (KR) provides the least information and structure while knowledge of correct response plus additional information. (KCR + I) provides the most information and explicit structure. Unfortunately, due to the numerous ways of combining information into feedback, the spectrum might not be so strictly defined. For example, it was possible to add additional abstract information to KR and only hint at the correct response. It was also possible to add an explicit explanation as to *why* the learner's response was incorrect (also (I)). In other words, the notion of a spectrum of content was viable, but each potential content type would have to be evaluated as to amount of information contained and explicitness of that information.

Typically, feedback content types beyond KR were not examined in the psychomotor literature. This was likely due to the difficult nature of controlling for content in feedback and the rigorous tradition of psychomotor skill research. Little focus occurred on variables that could not be tightly manipulated. This was not true of the educational and instructional literatures. Because of a desire to give feedback for complex cognitive tasks, they explored many forms of feedback content. KR, instead of being the focus, was generally assumed to be part of more complex feedback, such as knowledge of correct response (KCR) or KCR + information (KCR + I). The educational and instructional literatures focused on tailoring feedback content to the task to be learned, and at times the learner characteristics. This was due to their desire to develop effective instruction rather than precisely understanding *how* feedback affected human learning (Gagne, 1970). For example, KR seemed to match the learning outcome of declarative

knowledge and verbal learning tasks whereas KCR + I matched learning for intellectual or procedural skills (Smith & Ragan, 1993). Indeed, Smith and Ragan made separate feedback prescriptions for learning concepts, rules, procedural skills, problem solving, and psychomotor skills. For example, for *problem-solving* tasks, feedback should include KR plus hints, guidance, and/or error information. Once the learner achieves more expertise in the specific domain, feedback should include KR plus information on the speed or efficiency of their solution.

These conclusions by Smith and Ragan (1993) served an illustrative purpose. Minutely detailed prescriptions are seductive to designers, even if these data come from a single study or domain. This was exactly why we needed an over-arching understanding of what made feedback effective, else only overly specific prescriptions would exist, with little or no way to predict effectiveness in new or emerging situations. Feedback content could be designed to help learners achieve (McKendree, 1990; VanLehn, 1996), but we still needed principles to follow. Controlled empirical and observational studies provided evidence on which to base these principles.

One recent study compared KR to varying degrees of KR + I and KCR (Goodman et al., 2004) in learners making managerial decisions. This study included transfer and retention measures and found that as feedback content increased, so did performance during acquisition. However, actual learning declined. Also, more feedback resulted in less exploration of incorrect responses and arguably less depth of processing and self-assessment (Goodman et al., 2004).

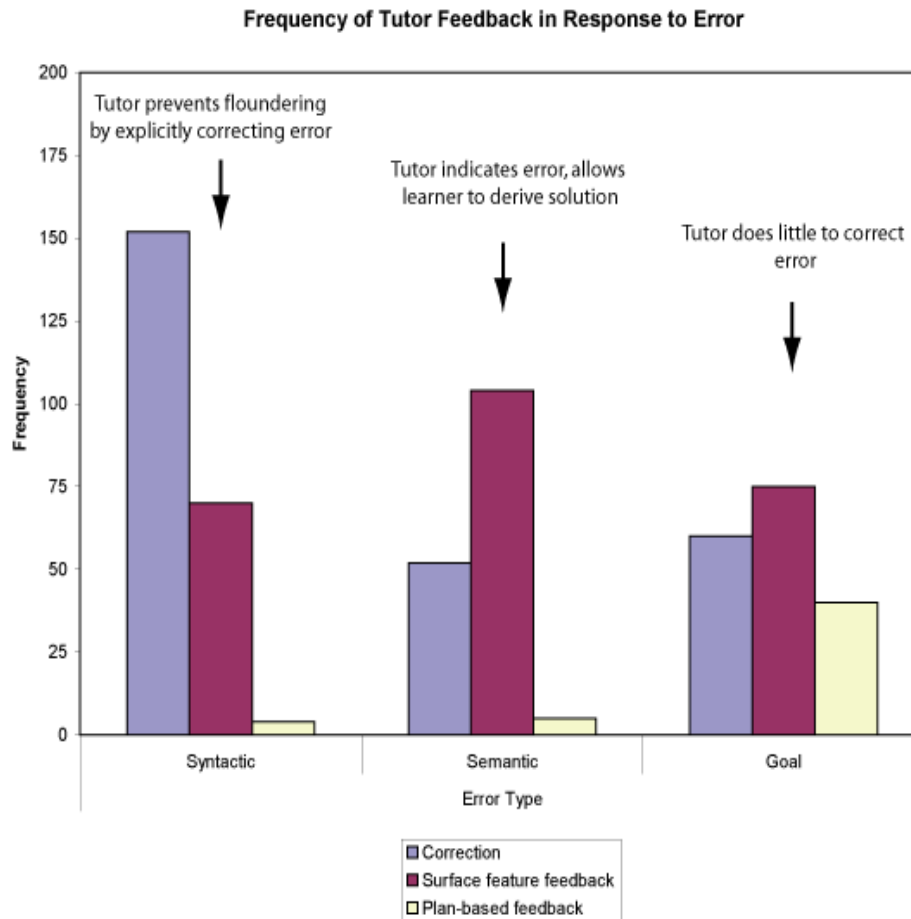
Other research showed that learners only developed higher-level problem solving strategies when they lost access to immediate feedback. In Cope and Simmons (1994), learners solved problems on a screen using a “turtle” that pointed in the next direction it moved. Learners who did not get to see the turtle point demonstrated greater learning of the task and higher-level strategies. Those who saw the turtle did not demonstrate higher-level problem solving strategies.



Human tutors appeared to tune feedback content to promote internal learner feedback. In one study, students brought a difficult classroom topic to a tutor and the tutoring conversations were analyzed (Fox, 1991). As a learner worked through a problem, human tutors gave confirmatory feedback to each *correct step* and each *incorrect* step that differed. In this way, tutors gave KR as feedback, the minimum amount possible, and did so immediately (Fox, 1991). The tutors indicated an incorrect action immediately as well, but allowed the learner to internally assess the error. However, human tutors did not offer confirmatory feedback after a correct action in all studies; some offered feedback mainly on errors (Merrill et al., 1995) or offer feedback that is brief, verbal, and positive or negative based on intonation (Graesser et al. 1995).

With respect to feedback content after an error, there were mixed findings within the literature. Some posited that having to recover from an error inhibited learning and prescribed the learner stay on the correct path as much as possible, to the point of being held to an ideal learner model of the task (Anderson, 1983; Lewis & Anderson, 1985; Sweller, 1988). However, other research pointed in the opposite direction; recovering from errors and coming to an impasse was the only way to trigger learning (Chi et al., 1989; Collins & Brown, 1988; Laird et al., 1986; Schank & Leake, 1989; VanLehn, 1990; Sleeman & Brown, 1982).

Human tutors provided different feedback content based on error type (Merrill et al., 1995): *correction feedback* (KCR), “That ‘a’ should be a ‘b’.”; *surface feature feedback*, KR about which feature was misused and direction to where the error was on the screen; and *plan based feedback*, reiterating the original goal of the current task with no other information. This was essentially KR, because it indicated to the learner he or she had pursued the wrong goal but offered no guidance (Figure 5).



*Figure 5.* Chart adapted from Merrill et al. (1995). Frequency of tutor correction is on the y-axis and the categories of error type are shown on the x-axis. Tutors were most likely to explicitly correct syntactic errors, point out semantic errors, and simply restate the plan for goal errors. This demonstrates how tutors tend to respond to error types with differing feedback content.

Elaborating, or adding more information to the feedback can result in a higher percentage of errors corrected on a post-test (Balzar, Doherty, & O'Connor, 1989; Roper, 1977). As more effort was expended to understand the feedback, the material was better learned, leading to improved retention test performance. These results were not always replicated (Hodes, 1984; Richards, 1989). An explanation for the division came from a study that found that elaborative feedback increased correct responses on a post-test with non-serious errors but not with serious or complex errors (Birenbaum & Tatsuoka, 1987). As more cognitive resources were available to

the learner (non-serious errors), more information could be absorbed as feedback, but as errors become serious and required more resources for the task, elaborative feedback became less effective.

*Summary.* Evidence such as this helps to develop general principles of feedback content, rather than minute prescriptions for each task. Having the amount and structure of content tuned to promote development of internal learner feedback resulted in more learning. Understanding content is resource-dependent. Content should be the least possible and never overwhelm the learner. Exactly how to determine what will overwhelm a particular learner is yet unknown.

### *Precision*

Feedback is a piece of information like any other and processed as such. Therefore, some said it was possible to prescribe a time length for feedback presentation based on complexity, precision, and length. Overly precise feedback increased complexity and therefore processing time of the feedback (Rogers, 1974). Self-paced presentation of feedback sped up across training time as humans learned to process the feedback just as they learned the task for which they received feedback. Newell proposed a numerical figure for KR presentation time in a simple task: 1 second (Barclay & Newell, 1980). These findings suggested that every skill contained the second skill of self-assessment via feedback within it, which also required processing.

Precision of the feedback was indicative of the corrective, informative nature of feedback rather than its reinforcement value. The higher the precision of the feedback, the more time needed to process the feedback, but this generally resulted in better learning to a point. Humans would not use feedback of “you were 2.113524337 inches too far” better than “you were 2.11 inches too far.” (Barclay & Newell, 1980; Rogers, 1974). Also, precise and prescriptive feedback

tended to mirror the findings of elaborative feedback content: desirable to a point, but too much precision resulted in less retention and transfer as learners used the feedback rather than developing their own self-assessment skills (Goodman et al., 2004).

*Summary.* Decisions about feedback precision should be made according to time of presentation of feedback, learner characteristics, and need for precision in the task. The findings for feedback precision conceptually match those for feedback content.

### *Timing*

The effect of delay of feedback has been studied in all skill domains. In psychomotor skill acquisition, the literature divided timing of feedback into several possible delay periods. Immediate feedback indicated no delay between response and feedback. Between each response was the inter-trial interval, that was subdivided into the pre-feedback interval (time between response and feedback) and post-feedback interval (time between feedback and next response).

Educational psychologists have numerous terms for feedback timing intervals (i.e., immediate, delayed, after item, after action review) which complicates interpretation of the research. For example, a meta-analysis (Kulik & Kulik, 1988) concluded “immediate” feedback was more beneficial than delayed feedback, but categorized the studies according to how the authors originally described the delay. Upon closer examination, feedback that was described as delayed was actually faster than in other studies where the feedback was immediate.

Generally, researchers argued that immediate feedback did not allow enough time for self-assessment and self-error correction, therefore harming the learning process (Kulhavy & Anderson, 1972; Kulhavy & Stock, 1989; Lewis & Anderson, 1985; Schroth, 1992). This was part of the belief that feedback was only present to guide the learner until he or she self-assessed

accurately (*cf.* Adam's subjective reinforcement, Salmoni's guidance/motivation). However, learners given control over when or whether they received feedback (and were forced to complete the task correctly before moving on) did not perform better than those given immediate feedback and forced to correct the error at that point (Corbett & Anderson, 1990). Careful examination of the task used in the Corbett and Anderson study (learning LISP) indicated there was little opportunity for self-assessment by the group that controlled their own feedback presentation, so the benefit of learner-controlled feedback remains a possibility if learners were given such an opportunity. However, providing immediate feedback for some tasks kept learners from floundering, and those learners achieved the same learning in less time than those allowed to flounder (Anderson, Corbett, Koedinger, & Pelletier, 1995). Human tutors also tended to offer rapid feedback or end a difficult task in order to prevent floundering at all costs (McArthur, Stasz, & Zmuidzinas, 1990). In fact, tutors did not allow much time for students to find errors; if the student did not say something almost immediately, the tutor intervened (Merrill et al., 1995).

A recent experiment compared immediate and delayed feedback for a category learning task and supposedly found an interaction of delay and task-type on learning (Maddox, Ashby, & Bohil, 2003). The first category learning task was rule-based, meaning that each category had an explicit, easily verbalizable rule for categorization of Gabor patches. The second task was an information integration task, where more accurate performance resulted from integrating perceived dimensions before coming to a categorization decision. Feedback timing was manipulated by giving KR after a .5 s (immediate), 2.5 s, 5 s, or 10 s (all delayed) mask. The general results were that delay had no effect on categorization accuracy in the rule-based task. However, delaying feedback in the information integration task resulted in less accuracy.

There were several concerns about these conclusions. First, this experiment violated

Salmoni et al.'s rules for acquisition versus learning (1984). Feedback, whether immediate or delayed, was present throughout the experiment and no tests of retention or transfer were conducted. Second, the immediate feedback group was given a longer post-feedback interval in which to potentially process both their initial response and the feedback itself. Also, the information integration group was likely unable to keep the image of the stimulus through the mask and could not encode it in words as the rule-based group did. This kept them from any benefit of self-reflection on the feedback received. Last, the authors admit information integration was a more complex task than rule-based categorization (Maddox et al., 2003). As per the discussion of feedback precision in this report, complexity translated to a need for longer feedback study time that the delayed group did not receive. In short, the need for self-assessment and potential benefits of delayed feedback are not negated by these findings.

*Summary.* Delay should be long enough to permit self-assessment of performance in acquisition before feedback is presented. Length of delay might not be important as long as it was sufficient for processing, but too long of a delay will require re-presentation of the stimulus.

### ***Trials-delay***

Trials-delay refers to feedback presented for a trial after subsequent trials had passed. Providing trials-delayed feedback increases learning. Lavery (1964) found that performance during acquisition was slower for a group who received 1-trial-delayed feedback than the 0-delay group but on a retention test with feedback removed, the 1-trial-delayed group outperformed the 0-delay group; this benefit lasted for three days of retention testing. In fact, the learning curves for the condition with 1-trial delay of feedback looked as though they were still receiving feedback on the retention tests, as they continued to improve their performance (Lavery, 1964).

However, trials-delayed feedback has not been extensively studied, and there were no articles investigating it later than 1979. This is likely due to research on summary feedback showing more promise. In fact, summary feedback is sometimes called trials-delayed feedback (Salmoni, et al., 1984). Providing summary information after a series of trials aided psychomotor skill acquisition (Gillespie, 2003; Wishart et al., 2002) and promoted learning more than occasional feedback on a single trial (Gillespie, 2003; Schmidt & Bjork, 1992; Schmidt & Wulf, 1997; Schooler & Anderson, 1990). Learners were presumably thinking back over their performance and applied the summary feedback. On trials without summary feedback, learners needed to assess their own performance. This evidence supported the idea that less frequent feedback encouraged the development of self-assessment behavior.

*Summary.* As with other feedback variables, trials-delayed feedback that promoted self-assessment between trials resulted in better learning.

### ***Frequency***

Feedback frequency, sometimes called scheduling, is typically divided into two subtypes: absolute feedback which is the number of times feedback was presented, and relative feedback which is the percentage of trials on which feedback was presented.

### ***Absolute***

Most theories of psychomotor learning predicted that more absolute feedback would lead to increased skill acquisition (Adams, 1987; Schmidt, 1975). Frequent feedback was typically beneficial (Anderson et al., 1995; Fox, 1991; Merrill et al., 1995) and the benefits were evident both during acquisition and on retention or retrieval tests (see Salmoni et al., 1984 for a review).

There was occasional evidence to the contrary, where lower absolute frequency resulted in the same performance during acquisition (e.g., Winstein & Schmidt, 1990).

### ***Relative***

Relative frequency of feedback seemed to have mixed effects on learning. Feedback given on every trial tended to produce less learning than feedback given on a sparser schedule, unlike absolute feedback. Reduced frequency feedback showed entirely different effects in acquisition versus learning. In acquisition, with feedback present on a certain number of trials, a higher percentage of feedback trials equated with better performance (Schroth, 1997). However, the opposite was true for trials demonstrating learning after feedback was removed. For learning, a *decreased* percentage of trials with feedback equated to better performance on a retention or transfer test (see Salmoni et al., 1984, for a review; Schroth, 1997).

It could be argued that absolute feedback frequency must be held constant to compare two groups with differing relative frequency. The problem was that holding absolute feedback constant automatically increased the number of total trials (and practice) of the group that received a lower relative feedback frequency. As an example, if absolute feedback were held to 25 trials with feedback, a low relative frequency group needed to perform 100 practice trials, 25 of which received feedback, resulting in a 25% presentation of KR. A high relative feedback group might receive only 25 practice trials to hold relative feedback at 100%. This was an obvious confound and could be used to explain the better retention or transfer performance with low relative frequency: more practice. There is evidence that reduced relative feedback aided learning even when absolute amount of feedback was reduced (e.g., Proteau & Isabelle, 2002; Schmidt & Bjork, 1992; Schooler & Anderson, 1990; Schroth, 1992).



There was some evidence that as task complexity increased, so did the need for higher relative feedback. In learning to slalom ski on a simulator, groups given 100% feedback on their actions outperformed a 50% and 0% condition on a retention test (Wulf, Shea, & Matschiner, 1998). The conclusion from this study should not be that less feedback is generally better; it should be that learning will be optimized when the appropriate feedback is provided for the task and for complex tasks more feedback may be required.

One argument against reduced feedback as mechanism for increased self-assessment was that it could be said better performance in transfer and retention was due to the similarity between training and test. In other words, less feedback during acquisition was more similar to the no-feedback retention test than high feedback during acquisition. At least one study provided evidence this was not the case. When given a delayed retention test with feedback included, the group who received less relative feedback in acquisition still outperformed the 100% feedback condition (Winstein & Schmidt, 1990). In another study, complex intellectual skills (such as generalizing analogies), were found to transfer least when participants were given feedback on every trial (Phye, 1989).

*Summary.* Enough feedback was needed to maintain somewhat accurate performance in acquisition, however too frequent feedback translated into a crutch for the learner. With such a crutch present, learners tended not to engage in the other (perhaps more important) cognitive activities such as self-correction and self-assessment

### ***Interpolated Activities***

An interpolated activity occurred in the inter-trial interval during skill acquisition. This activity might be related or unrelated to the skill being learned and occur before or after feedback

was presented. In acquisition, resource demanding interpolated activities degraded performance (Shea & Upton, 1976). Such activities might interfere with memory for prior responses and leading to less corrective planning on the next response (Salmoni et al., 1984). This view was echoed in the interference perseveration hypothesis (Kulhavy & Anderson, 1972), except that the interpolated activity was predicted to interfere with the learner's memory of the initial, incorrect response. However, these findings addressed the data for studies of performance during acquisition. When learning was examined, interpolated activities showed the same effects *unless the activity was to estimate the correctness of the prior response* (Guadagnoli & Kohl, 2001; Hogan & Yanowitz, 1978). One of the more recent and well-executed studies of interpolated activities was carried out by Swinnen in 1990. In a series of three experiments, Swinnen manipulated the presence of an interpolated activity either in the pre or post-feedback interval between trials of learning a timed movement. The absolute inter-trial interval was kept the same across conditions. Off-task interpolated activities increased errors both in acquisition and retention, while errors were reduced by providing no activities or asking the learner to estimate the correctness of the last trial (Swinnen, 1990). Other studies also found estimating correctness as an interpolated activity facilitated learning. It was theorized this was due to forcing in-depth processing of the task and the skill of self-assessment (Lhyle & Kulhavy, 1987). Even giving the option for more in-depth processing or self-assessment could improve learning prior to any elaboration past feedback (Richards, 1989).

*Summary.* Interpolated activities that distract from self-assessment before and after receiving feedback inhibit learning. Interpolated activities promoting self-assessment promote learning.

## Chapter 4 – Learner Characteristics

*“The status of a person at the time influences his response...Equally obvious are the influences of those more permanent attitudes or sets of mind in a person which belong to him as...optimist or pessimist, realist or romanticist, extravert or introvert, hard-boiled or sensitive, energetic or inert.”*

-Thorndike, 1931a, p. 120)

Assuming an interaction of feedback and individual differences was not untoward; links between person variables and learning stretched back to the beginnings of psychological research. In the literature, characteristics that have been studied include immutable learner characteristics such as mental resources, abilities, cognitive aptitudes, and working memory capacity, and mutable learner characteristics such as experience, motivation, and arousal levels.

### *Immutable Learner Characteristics*

Immutable learner characteristics refer to those that are relatively stable over time and not easily changed by experience. They are the capabilities that the learner brings to the task.

Verbal protocols of complex cognitive skill acquisition indicated that most learners self-assessed when provided with feedback they were in error. However, poor students (those who performed badly on subsequent retention and transfer tests) tended to respond positively that they now understood the concept. Good students responded with requests for clarification or an analysis of why they performed incorrectly (Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Pirolli & Recker, 1994). The incompetent lacked the ability to realize their incompetence; thus, *accuracy* in self-assessment was an important component of learning. Incorrect self-monitoring lead to misapplication of effort and harmed learning (Kanfer & Ackerman, 1989). In other words, self-assessment through feedback could contribute to learning, but learner had to have the capacity to take advantage of that feedback.

### ***Cognitive Aptitudes***

Educational psychologists referred to abilities as cognitive aptitudes. Snow (1997) developed a taxonomy of cognitive aptitudes that specifically contributed to learning. The taxonomy include fluid intelligence, crystallized intelligence, idea production, and memory. For perceptual speed, the taxonomy included visual-spatial perception, auditory perception, and general speediness. Although this taxonomy was specific to cognitive skill, the perceptual motor skill of reaction time was included as well. Others have also argued that there are three main classes of ability important for learning: general intelligence, perceptual speed, and psychomotor ability (Kanfer & Ackerman, 1989).

### ***Working Memory Capacity***

Working memory capacity (WMC) is defined as an “a system of: (a) short-term “stores,” consisting of LTM traces in a variety of representational formats active above a threshold; (b) rehearsal processes and strategies for achieving and maintaining that activation; and (c) executive attention.” (Engle & Kane, 2004, p. 146; Kane & Engle, 2003). WMC also highly correlates with language comprehension, reasoning ability, and general intelligence (Engle & Kane; Woltz, 1985).

Because WMC is related to the ability to hold and integrate information in the memory store, WMC should relate to task complexity (Brumback, et al 2005). WMC highly correlates to reasoning ability (Woltz, 1985), and if understanding and utilizing feedback is actually part of the task, higher reasoning ability should result in more understanding of feedback provided. Additionally, WMC can be used to explain older adult’s problems with certain tasks (Light & Anderson, 1985). It is not age, per se, that causes trouble with reasoning tasks, it is a decrease in

WMC. Additionally, WMC is not just correlated with reasoning ability and Spearman's  $g$ , it's correlated with attentional control (Feldman-Barrett, Tugade, & Engle, 2004) and therefore information integration (Heitz, Unsworth, & Engle, 2004).

WMC provides an index of how well an individual will perform on many types of tasks (Turner & Engle, 1989). That is, measuring an individual's WMC predicts how that person will learn a new task, when things such as experience, arousal, and motivation are held equal. Most practically, WMC is a good choice from which to predict effective feedback for a task because it can be assessed fairly quickly and easily (Kane & Engle, 2004) and is a fairly enduring characteristic.

*Summary.* Choice of nearly every feedback variable depends on learner ability. Feedback should give the learners what their ability level can handle, which will be different for differing abilities.

### ***Mutable Characteristics***

Mutable learner characteristics refer to those that change over time through experience or education. They also include the state of the individual at different points in time (e.g., arousal level or motivation).

### ***Experience***

*"Except perhaps in infancy, a great part of the variance in rates of learning between individuals must be attributable to differences in the results of past learning."*

-Estes, 1970, p. 31

According to the closed-loop theory of motor learning, the novice learner specifically requires feedback whereas the expert might not due to possessing internal subjective reinforcement. This implied a gradual lessening of feedback as the learner gained expertise, but

did not demand it. In general, less feedback was beneficial to learners closer to the level of expert. The less prior knowledge learners had or brought to a task, the more elaboration they typically needed, however also the less elaboration they were likely able to process (Sweller, 1988). The question of what feedback to provide based on experience was therefore likely a function of the task, but not an easy one to answer given the paradox just stated.

Prior experience was a predictor of future performance on a task. However, this is also an open question as there have been mixed findings of what may be used to predict intermediate levels of performance. Some models stated that abilities are predictive above and beyond prior performance (Ackerman & Cianciolo, 2000) but in other cases prior performance was the only significant predictor (Czaja, Sharit, Ownby, & Roth, 2001). It was likely the best predictor depended on the demands of the task, such as whether it was consistent or varied (Farrell & McDaniel, 2001; Rogers, Fisk, & Hertzog, 1994). More research in this area will be useful to understand whether to base feedback on ability, prior experience, or some weighting of the two.

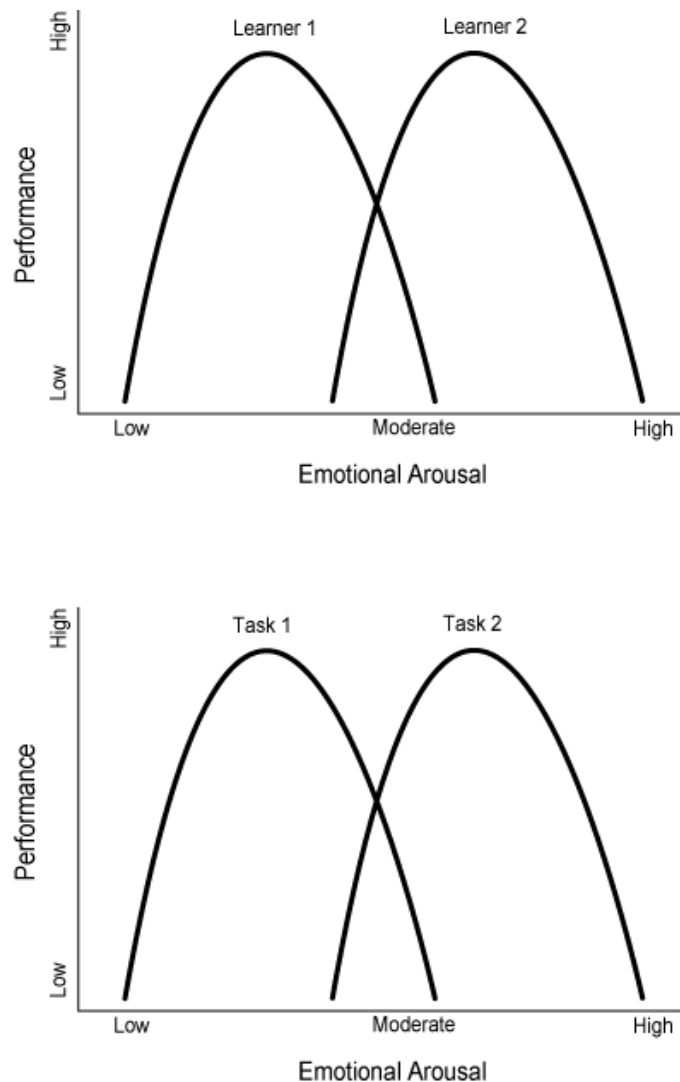
*Summary.* According to their experience level, learners may receive as much feedback as they can process, however it is better to provide the least feedback possible that aids the learner in filling in the rest of what they can process themselves. For example, if a learner could process elaborative feedback on a conceptual component of the task, this may be provided, but it would be better to provide feedback that allows the learner to internally develop the concept (if this is possible). Further, as prior experience predicted future experience, this variable might be used to set feedback in a task once the learner had some experience.

### ***Arousal***

*“The greater the satisfaction or discomfort, the greater the strengthening or weakening of the bond.”*

-Thorndike, 1911, p. 244

A commonly mentioned psychological concept is the Yerkes-Dodson law (Yerkes & Dodson, 1908). The law stated if a task were complex or required fine motor skill, the optimal level of learner arousal was low. If the task were relatively simple, requiring gross motor skill, the optimal level of arousal was high (Yerkes & Dodson, 1908). In their own words, “a stimulus



*Figure 6.* Illustration of how the Yerkes-Dodson law might differ for individual learners or tasks. The y-axis shows performance level and the x-axis emotional arousal. Optimal arousal levels for performance may depend on learner characteristics and task demands

whose strength is nearer to the threshold than to the point of harmful stimulation is most favorable to the acquisition of a habit” (Yerkes & Dodson, 1908, p. 481). Yerkes and Dodson

stated the optimal level of arousal was lower for more difficult tasks (the learners needed to concentrate on the material), and higher for tasks that required persistence (the learners needed more motivation). A recent study where arousal was induced via caffeine supported these principles (Watters, Martin, & Schreter, 1997).

The allure of the Yerkes-Dodson law was that it appeared to make sense; there seemed to be some relationship between arousal and performance.

However some studies found arousal and performance to have other shaped functions (Hancock & Ganey, 2003; Matthews, Davies, & Lees, 1990). Some researchers tackled this diverging evidence by dividing arousal into subdivisions, such as wakefulness and vigor. Wakefulness did not follow the inverted U function, but vigor did (Dickman, 2002), which might help to disentangle the mixed findings. It was also possible effects of arousal differed by learner experience and by task (Figure 6), increasing the complexity of pinpointing optimal arousal levels.

*Summary.* It is not yet understood exactly the role arousal plays in learning. However, very high arousal may result in reduced resources for processing feedback or self-assessing performance during acquisition, reducing learning. Different subcategories of arousal must be assessed for the data to reveal consistent patterns.

### ***Motivation***

*“Motivation refers to the direction of attentional effort, the proportion of total attentional effort directed to the task (intensity), and the extent to which attentional effort toward the task is maintained over time.”*

-Kanfer & Ackerman, 1989, p. 661

Motivation might occur before beginning at task (“distal”) which is outside the task and



feedback cycle. Motivation during task acquisition (“proximal”) is more relevant to issues of feedback because occurs during the task. Proximal motivation consists of self-monitoring, evaluation and reaction and “may compete with on-task...demands” (Kanfer & Ackerman, 1989, p. 662).

In studies where the answers were available before students were required to make a response, they tended to copy the answers rather than exert effort (Anderson, Kulhavy, & Andre, 1971, 1972). When allowed to control the feedback they received, learners used it as crutch. Poor students especially tended to consult feedback and examples when given the opportunity, rather than self-evaluate (Chi et al., 1989). This was an example of a low-motivation behavior. However, setting performance goals motivated learners (Kanfer & Ackerman, 1989). Feedback that encouraged goal-setting and accurate self-assessment should result in better learning. Though as will be discussed later, overall goal setting may be more appropriate for experienced or high-ability students.

*Summary.* Feedback should be designed both to motivate and respond to the motivation levels of the student. For example, an unmotivated student will likely not put forth resources to perform without feedback, but also will use feedback as a crutch if supplied too readily.

Mutable learner characteristics may be measured or assessed before learning. However, unlike immutable characteristics, they may actually change during acquisition. Feedback should optimally be sensitive to changes and adapt to the appropriate level of feedback. For example, as learners gain experience, they could likely comprehend more conceptually based feedback.

## Chapter 5 – Task Demands

*“Situations with low processing demands benefit from practice conditions that increase the load and challenge the performer, whereas practice conditions that result in extremely high load should benefit from conditions that reduce the load to more manageable levels.”*

-Wulf & Shea, 2002, p 185

Accuracy of self-assessment may not only be affected by the ability of the learner, but also by the demands of the task. Self-assessment may be manipulated by changing task demands. For example, a way to improve accuracy in self-assessment was to force self-generation of feedback (Swindell & Walls, 1993). Of course, depending on learner characteristics, not all learners may profit from that technique. It has even been suggested that each learner has a “challenge point,” or an optimal point at which the correct amount of informative feedback should be provided for a certain level of task difficulty (Guadagnoli & Lee, 2004). Specific guidelines were not outlined, but this appears to be an interesting area for future research.

### *Complexity*

One way to describe task demands is in terms of complexity. Increasing complexity requires increasing the number of cognitive stages required to perform a task (Sternberg, 1969). For example, adding components to problems whose answers interact and depend on one another increases complexity.

If a task does not require much processing (simple task), an increased load during practice results in better learning, whereas for a task that requires heavy processing (complex task), practice that reduces the load on the learner seems to result in more learning (Wulf & Shea, 2002). This idea of practice as a “load” extends to feedback. Giving feedback on a simple task should result in the most learning when it contains more information or otherwise alters

content and timing to most challenge the learner. However, giving feedback on a complex task would require just the opposite; the load on the learner from the task is already high and the feedback needs to reduce that load.

### ***Practice Schedules***

Learners in blocked practice conditions greatly over-estimate their judgment of how well they learned the task (Simon & Bjork, 2001). They did this despite consistent findings that they tend to perform better in acquisition and worse in retention and transfer tests than learners in random practice conditions. Those in random conditions also tended to overestimate their anticipated retention performance, however, their judgments of their own learning were not as overconfident as those receiving blocked practice (Simon & Bjork, 2001). Therefore, accuracy in self-assessment may not only be resource driven but also task driven: learners in the easier blocked acquisition trials *should* have more free resources to judge their learning accurately, but this was not the case. This is why the structure of the task demands need to be carefully considered when providing feedback.

*Summary.* Optimal feedback is affected by task demands. For example, if the task demands included blocked trials, feedback is needed to counter the learner's over-confidence in retention of the skill. As task complexity increases, the need for feedback also increases to enable the learner to perform the task.

## Chapter 6 – Known Interactions

Arguably, all the variables discussed thus far can interact with each other at some level. However, certain interactions have empirical support. These included the interaction of learner characteristics and feedback content (aptitude-treatment interaction), frequency and timing of feedback, learner ability level and motivation, and feedback content and error type. Other interactions might prove to be just as important (such as motivation level and feedback timing), but those have not received empirical attention.

### *Aptitude x Treatment Interactions*

Because it seemed intuitive that learners of different skill levels would benefit from different types of feedback, researchers investigated the interaction of feedback variables and learner ability. This type of aptitude-treatment interaction prescribed different training for learners of differing experience or ability. If the characteristics of the learner and the type of training chosen for the task were well matched, the most learning occurred (Cronbach & Snow, 1977; Snow & Swanson, 1992). Individual differences such as general intelligence, reasoning, spatial, perceptual speed, and psychomotor abilities could determine complex skill acquisition (Ackerman, 1988, 1992; Ackerman & Cianciolo, 2000).

Despite interest in the interaction of feedback content and student skill level, there had been little empirical research investigating student skill level and appropriate feedback. Most studies compared one type of feedback to another for students of a particular skill level, which as discussed, was of limited usefulness without organizing principles or a model.

More structured training resulted in more learning from lower ability students, whereas less structured training resulted in more learning by high-ability students (Cronbach & Snow,

1977; Kozma, 1991; Tuovinen & Sweller, 1999). On multiple-choice tests, high-ability students learned more with the feedback type “answer until correct (AUC)” than “knowledge of correct response (KCR)” (Clariana & Smith, 1989). AUC was less structured and informative than KCR, and the high-ability students seemed to use the opportunity to reassess their answer. Lower ability students improved their scores more via knowledge of correct response feedback, which provided more structure for learning than AUC. AUC offered the chance to reflect on why an initial response was incorrect and reconsider other choices; however it also offered the chance to choose other responses without reflection until happening upon the correct answer. Low-ability students quickly selected other options until happening upon the correct response. In a study comparing knowledge-of-correct-response (KCR) and AUC feedback for low-ability students, KCR produced superior post-test performance (Clariana, 1990). However, in a later study with low-ability learners, KCR was found to be no different from AUC. As test questions required more inference and became more dissimilar from the pre-test material, AUC did show a slight advantage; the authors ascribed this result to the additional processing of correct and incorrect responses that occurred during AUC (Clariana, Ross, & Morrison, 1992). More structure for low-ability students did not necessarily result in as much as the less structured high-ability students but only in more learning than their low-ability counterparts (Clariana & Smith, 1989).

However, ability level was not the only significant learner characteristic; anxiety level was studied as well. Highly anxious students (categorized via measures of test anxiety) tended to learn more from structured training whereas non-anxious students tended to learn more from less structured environments (Snow, 1997).

Human tutors tend to adapt feedback to aptitude. Tutors of remedial students were found to be directive and suggested possible solutions as well as pointing out errors to students

(McArthur et al., 1990). However, tutors of students not classified as low-ability tended to avoid telling students why they were incorrect. They instead allowed, with some direction, students to discover their own errors (Fox, 1991; Merrill et al., 1995). Human tutors were dynamic feedback systems who chose tasks, evaluated learner knowledge, and corrected learner errors (McArthur et al., 1990). They did all of this dynamically in light of learner characteristics that they evaluated, such as amount of experience in the task and general ability level. For example, most tutors in the McArthur et al. study faded the frequency of their feedback as learners gained more task experience. They did this even when learners made errors in the task; they corrected the error but did not offer feedback on subsequent trials after the error. Interestingly, the tutors rarely explicitly assessed learner's prior knowledge but responded only to the error. The authors ascribed this behavior to the desire of the tutors to proceed quickly through the material rather than apply the most effective tutoring process, which seemed to apply to many human tutors. Tutors were not told how to teach, yet they were similar to each other and to other studies of tutoring (Merrill et al., 1995). Experienced tutors shared certain behaviors, though those behaviors did not always promote the most learning. As mentioned previously, deep processing of the material via feedback from the tutor (or any system) should result in more learning by the student. However, tutors did not apply advanced methods to encourage deep learner processing.

Though many studies assesses low-ability learners (e.g., Clariana, 1990), ability level was never systematically investigated. Ability level was included in a re-analysis of several dissertation studies and low prior knowledge learners did benefit more from KCR while high prior knowledge learners benefited from AUC (Clariana, 1993). The author hypothesized that low-ability learners were not self-reflective enough to use the information contained in AUC feedback whereas high-ability learners could and did use that extra information to add to their

processing of the material. However, these results came from one paper where the notion of “ability” and “prior knowledge” was used interchangeably. It was possible that these results might have been true for one group or the other but not both. The term “ability” is sometimes used vaguely in the educational literature, meaning either something measured by a particular ability test or simply equating remedial students with “low ability.” This could be an oversimplification, as according to Gardner’s multiple intelligence theory, learners may be low on some abilities while high on others (Gardner, 2003). A general label of low or high ability in a study was not a sufficient definition from which to draw conclusions.

There was, perhaps, a cognitive construct behind both of these findings. If higher ability students had more resources or a larger amount of mental resources, it explained why they were better able to learn in a less structured environment; they structured the training themselves rather than having it done for them. Similarly, anxiety consumed resources, equating those who were highly anxious to the resources of the low-ability learner. One expected the most resource-reduced group to be highly anxious, low-ability students, and indeed these were the lowest performers (Snow, 1989a, as cited in Snow, 1997).

*Summary.* More structured feedback would benefit lower ability and/or highly anxious learners whereas less structured feedback would benefit higher ability and/or less anxious learners.

### ***Timing x Error Type Interactions***

Human tutors base timing of feedback on error criticality. A fatal or critical error is one that ended progression through a task. Human tutors tended to intervene immediately when such errors arose as well as when an error distracted the learner from the goal and caused floundering

(Merrill, Reiser, Ranney, & Trafton, 1992). However, human tutors waited to give feedback on tactical or decision-making types of errors until students had completed enough of the task for discussion of those errors to be productive (Katz, O'Donnell, & Kay, 2000; Merrill et al., 1992).

Non-critical errors may not mandate immediate feedback (Anderson & Lebiere, 1998). After non-fatal errors, the student may be allowed to proceed uninterrupted through the task and receive notification of the error after the task was complete. An overload of feedback during the task can confuse and overload the learner (Schooler & Anderson, 1990; Sweller & Chandler, 1994).

*Summary.* Delaying feedback on non-critical errors reduced the cognitive load during task acquisition while maintaining the benefit of more elaborative feedback.

### ***Ability x Motivation Interactions***

Both high and low-ability learners benefited from increased motivation, but low-ability learners benefited the most (Kanfer & Ackerman, 1989). As was previously discussed, cognitive resources appeared to play a role in this interaction. For example, novices showed less ability to self-assess, likely because their learning stage was the most cognitively demanding. Specific to ability and motivation, lower motivation learners were also less likely to self-assess. Feedback as motivation may be most important for low-ability learners where students need to be able to assign failure to the difficulty of the problem rather than their own ability (Hoska, 1993). Such false ways of increasing self-efficacy were actually adaptive, as they could encourage the learner to progress with the task (committing the fundamental attribution error). Because humans are more likely to use heuristics under high cognitive load (Wigboldus et al., 2004), it is not surprising that it was the low-ability learners who fell prey to the attribution error.



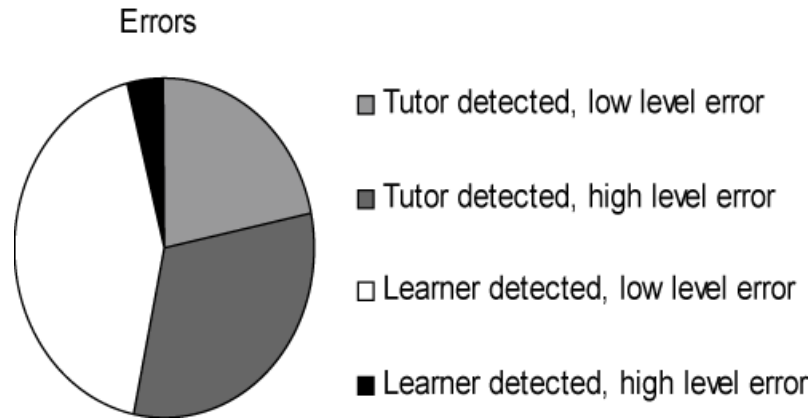
### *Content x Error Type Interactions*

The interaction of feedback content and error type was not one that had been systematically tested. Error type was often a post-hoc categorization, used to explain mixed findings. However, as with other feedback variables and interactions, it did make sense that error type would interact strongly with feedback content due to the resource-dependent nature of both.

In a study of learning to program in LISP, human tutors gave differing feedback content based on error type (Merrill et al., 1995). Tutors offered less and less explicit direction. Feedback for syntactic error was quick and explicit. Feedback on semantic errors was less explicit, but still directed the learner. Goal errors prompted a re-statement of the plan, the least directive form of feedback. The actions of the human tutors was conceptually similar to the elaboration on non-serious errors in the Chi et al. study (1989); learners were given feedback by tutors according to what their resources could handle for a particular type of error.

There was disagreement about errors in general. Some suggested having to recover from errors was detrimental (Anderson, 1983; Lewis & Anderson, 1985; Sweller, 1988), but others claimed recovery and explanation of impasses was critical (Chi et al., 1989; Laird et al., 1986; Schank & Leake, 1989; VanLehn, 1990). Human tutors allowed students to do as much as possible when learning could progress through error self-recovery. But if costs of error repair were high, tutors simply told students what to do and kept them on track (McArthur et al., 1990).

It is likely that learning from error depended on the task and the seriousness of the error. Human tutors did not leave errors unnoticed or unmentioned for long; they detected and repaired them rapidly (Fox, 1991). In the Merrill et al. (1995) study, half the errors were caught by the learner and most of these were low level errors without serious consequences (Figure 7). Tutors caught and corrected a much larger amount of high-level errors than low-level ones.



*Figure 7.* Pie-chart developed from Merrill et al, (1995) showing who tended to correct errors during tutoring. Learners and tutors each corrected about half the total errors during a lesson, but tutors detected and corrected more high-level errors than did students.

As mentioned previously, when feedback content was more elaborative and errors were not serious, post-test performance increased. However, when errors were serious or complex, additional elaboration resulted in less learning demonstrated at post-test (Birenbaum & Tatsuoka, 1987). Again, this interaction appeared to depend on learner resources where an overload results in less learning (Sweller, 1988).

*Summary.* As more resources were available, elaborative information aided learning, but for more serious errors and when more resources were required for the *task*, elaborative feedback became less effective as it overloaded the learner.

### ***Summary of Known Interactions***

To prescribe effective feedback, it is not enough to know how each feedback variable, learner characteristic, and task demand affects learning. Their interactions must be understood. However, the piecemeal approach of measuring interaction after interaction will at best produce a taxonomy for designers. The data point to common principles present in the main effects: effective feedback needs to accommodate learner characteristics and task demands. Extracting principles of feedback for application should apply both to main effects and interactions.

## Chapter 7 – Principles of Feedback Design

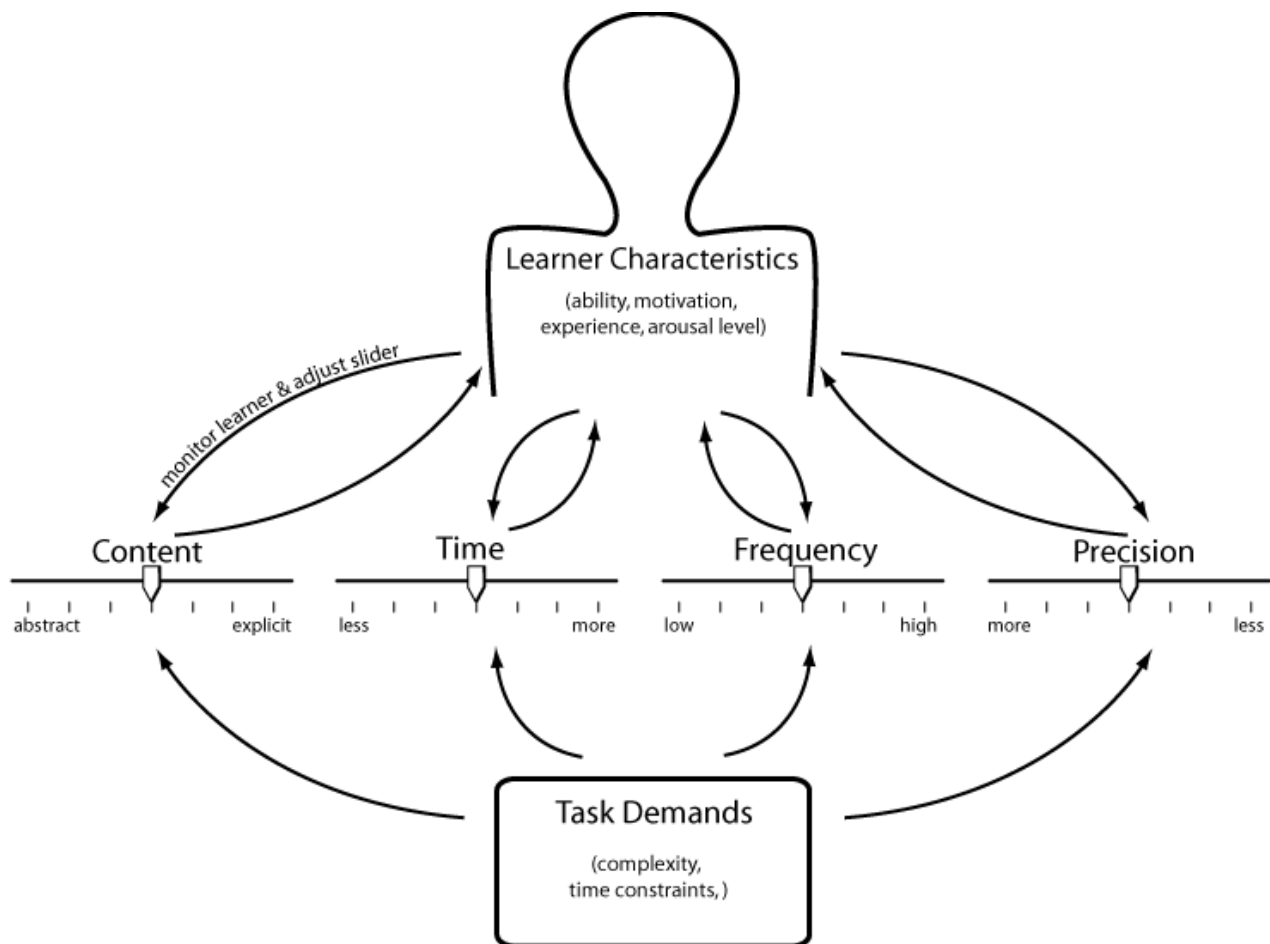
The first principle is that feedback must guide performance in acquisition until the learner is capable of generating internal feedback. In essence, the function of feedback is to *train* the skill of self-assessment in a task.

The second principle is that this guidance should be calibrated to the resources of the learner and demands of the task. Feedback should enable the learner to perform the task without but not become a crutch for performance. Such calibration is linked to the feedback variables and learner characteristics discussed in Chapters 3 and 4.

For example, general findings concerning aptitude-treatment interactions were common. Undeniably, the type of feedback must be calibrated to the ability or experience level of the learner. Researchers found a need for more structure and explicit feedback for low-ability learners and the opposite for high-ability learners. Another way of framing this was to say that the most effective feedback was the minimum possible to keep learners performing the task and able to self-assess. High-ability learners did this with less structure and less explicit information than low-ability learners. Aptitude-treatment interactions supported this claim. Response certainty, an interpolated activity, supported this because it was practice in the skill of self-assessment. This explained the differential findings for delay of feedback. Did delay contribute to the skill of self-assessment? If so, the delay was productive for learning. This explains the precision/overload findings, “source of feedback” findings, and “content” findings. Will conceptual feedback aid more in learning than directive feedback? Only when it offers deeper processing opportunities and learners can do that processing.

## Chapter 8 – A Conceptual Model of Effective Feedback

Based on our systematic and extensive review of the literature, we have developed a conceptual model to reflect the variables relevant to effective feedback design to support performance and learning. In the most general sense, external feedback needs to be designed to make itself obsolete. Figure 8 explains how a dynamic external feedback system trains learners to self-assess during a task.



*Figure 8.* A conceptual model of dynamic feedback. The three main components of the model are learner characteristics, feedback variables, and task demands. Feedback variables are affected by both learner characteristics and task demands. In a feedback loop, the learner characteristics are monitored during learning to detect whether the feedback variables need be changed.

In this model, learner characteristics influence the selection and design of feedback variables. For example, if a learner had high ability, the slider for feedback content would be set toward the abstract rather than the explicit. If a learner had low motivation, feedback might be used to keep the learner engaged in the task, in which case feedback would be given more frequently. Moreover, some learners may require more time to process feedback than others which might influence the time at which the feedback should be presented. Where the sliders are set for a task is dynamic, and depends on learner characteristics at any point in time. For example, the arrows return from the feedback variables to the learner characteristics to monitor how much experience with the task the learner gained, and adjust the feedback sliders accordingly during learning. Feedback is partially controlled by task demands. For example, if the recommendation based on learner characteristics was to provide elaborative content, a time-constrained task may make that impossible. To make design decisions even more difficult, the sliders for the feedback variables are interactive and hence influence each other as well. Increasing feedback frequency may mean automatically reducing content to avoid overloading the learner. Many of these interactions have been empirically documented (see Chapter 6), but detailed information concerning their importance and weight still needs investigation.

This model is complex, however this complexity does not negate its usefulness. All of the sub-components of the feedback decisions were based on empirical data, including the interactions between the variables. There is much work remaining to establish exact design guidelines, but even a model in this stage has applicability.

## Chapter 9 – Conclusion

This project combines the issue of understanding the role of the human operator (in collaboration with automated systems) with the interface technology issue of translating data into feedback. Automated systems can collect potentially unlimited amounts of data. A critical challenge is ensuring that those data are presented properly to the human when needed for appropriate management of, and learning about, the automated system. Variables such as the content and the timing of this feedback information must be understood. The goal of this report was to provide such understanding and to provide insight about when feedback supports learning, when it interferes with learning, and how optimality of feedback differs across tasks, contexts, and users.

Based on a systematic, integrative analysis of the research literature, we developed a conceptual model of the critical factors that affect feedback efficacy (see Figure 8). The model illustrates the importance of considering the separate and interactive effects of learner characteristics, task demands, and feedback variables such as content, timing, frequency, and precision. Moreover, we recommend two general feedback principles:

1. Feedback must *train* operators to self-assess their performance on the task; that is, they must ultimately be able to provide internal feedback.
2. Feedback should be calibrated to the resources of the learner and demands of the task.

Feedback must enable the learner to perform the task without becoming a crutch.

We are currently designing empirical studies to develop a process to optimize system-generated feedback to match the performance and learning needs of the operator. This research will lead to feedback principles that will clarify the role of the human operator and prescribe how to design feedback for that person, for a particular task, and in a particular context.

## References Cited

- Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General*, 117, 288-318.
- Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. *Journal of Applied Psychology*, 77(5), 598-614.
- Ackerman, P. L., & Cianciolo, A. T. (2000). Cognitive, perceptual-speed, and psychomotor determinants of individual differences during skill acquisition. *Journal of Experimental Psychology: Applied*, 6(4), 259-290.
- Adams, J. A. (1978). Theoretical issues for knowledge of results. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 229-240). New York: Academic Press.
- Adams, J. A. (1987). Historical review and appraisal of research on learning, retention, and transfer of human motor skills. *Psychological Bulletin*, 101, 41-74.
- Anderson, J. R. (1982). Acquisition of a cognitive skill. *Psychological Review*, 89(4), 369-406.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R. (1996). A simple theory of complex cognition. *American Psychologist*, 51(4), 355-365.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036-1060.
- Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4, 167-207.
- Anderson, J. R., & Lebiere, C. (1998). Newell's dream. In J. R. Anderson and C. Lebiere (Eds.), *The atomic components of thought* (pp. 1-18). Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, R. C., Kulhavy, R. W., & Andre, T. (1971). Feedback procedures in programmed

- instruction. *Journal of Educational Psychology*, 62, 148-156.
- Anderson, R. C., Kulhavy, R. W., & Andre, T. (1972). Conditions under which feedback facilitates learning from programmed lessons. *Journal of Educational Psychology*, 63, 186-188.
- Annett, J. (1989). Training skilled performance. In A. M. Colley and J. R. Beech (Eds.), *Acquisition and performance of cognitive skills* (pp. 61-84). Chichester: Wiley & Sons.
- Balzar, W. K., Doherty, M. E., & O'Connor, R. (1989). Effects of cognitive feedback on performance. *Psychological Bulletin*, 106, 410-433.
- Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *Journal of Abnormal and Social Psychology*, 63, 575-582.
- Barclay, C. R., & Newell, K. M. (1980). Children's processing of information in motor skill acquisition. *Journal of Experimental Child Psychology*, 30(1), 98-108.
- Bilodeau, E. A., & Bilodeau, I. M. (1958). Variable frequency knowledge of results and the learning of a simple skill. *Journal of Experimental Psychology*, 55, 379-383.
- Birenbaum, M., & Tatsuoka, K. K. (1987). Effects of "on-line" test feedback on the seriousness of subsequent errors. *Journal of Educational Measurement*, 24(2), 145-155.
- Brackbill, Y., Bravos, A., & Starr, R. H. (1962). Delay-improved retention of a difficult task. *Journal of Comparative and Physiological Psychology*, 55(6), 947-952.
- Buff, W. L., & Campbell, G. E. (2002). *What to do or what not to do? Identifying the content of effective feedback*. Proceedings of the Human Factors and Ergonomics Society 46<sup>th</sup> Annual Meeting, Baltimore, MD.
- Burton, J. K., Moore, D. M., & Magliaro, S. (2004). Behaviorism and instructional technology. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2<sup>nd</sup> ed., pp. 46-67). Mahwah: Erlbaum.
- Butler, D., & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65, 245-281.
- Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM*, 23(7), 396-410.



- Carlson, R. A., & Yaure, R. G. (1990). Practice schedules and the use of component skills in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 484-496.
- Carroll, W. R., & Bandura, A. (1982). The role of visual monitoring in observational learning of action patterns: Making the unobservable observable. *Journal of Motor Behavior*, 14(2), 153-167.
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145-182.
- Clariana, R. B. (1990). A comparison of answer until correct feedback and knowledge of correct response feedback under two conditions of contextualization. *Journal of Computer-Based Instruction*, 17(4), 125-129.
- Clariana, R. B. (1993). A review of multiple-try feedback in traditional and computer-based instruction. *Journal of Computer-Based Instruction*, 20(3), 67-74.
- Clariana, R. B. (1999). Differential memory effects for immediate and delayed feedback: A delta rule explanation of feedback timing effects, *Association for Educational Communications and Technology Annual Convention*. Houston, TX.
- Clariana, R. B., Ross, S. M., & Morrison, G. R. (1992). The effects of different feedback strategies using computer-administered multiple-choice questions as instruction. *Educational Technology Research and Design*, 39(2), 5-17.
- Clariana, R. B., & Smith, L. B. (November, 1989). *Comparative effects of ability and feedback form in computer-assisted instruction*. Presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR.
- Cope, P., & Simmons, M. (1994). Some effects of limited feedback on performance and problem-solving strategy in a logo microworld. *Journal of Educational Psychology*, 86(3), 368-379.
- Corbett, A. T., & Anderson, J. R. (1990). *The effect of feedback control on learning to program with the lisp tutor*. Paper presented at the Twelfth Annual Conference of the Cognitive Science Society, Cambridge, MA.

- Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction*, 9, 109-142.
- Craik, F. I. M. (2002). Levels of processing: Past, present, and future? *Memory*, 10(5), 305-318.
- Craik, F. I. M., & Lockhart, R. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- Czaja, S. J., Sharit, J., Ownby, R., & Roth, D. L. (2001). Examining age-differences in performance of a complex information search and retrieval task. *Psychology and Aging*, 16(4), 564-579.
- Deakin, J. M., & Proteau, L. (2000). The role of scheduling in learning through observation. *Journal of Motor Behavior*(32), 3.
- Dickman, S. J. (2002). Dimensions of arousal: Wakefulness and vigor. *Human Factors*, 44(3), 429-442.
- Dunning, D., Johnson, K., Ehrlinger, J., & Druger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12(3), 83-88.
- Elliott, D., & Jaeger, M. (1988). Practice and the visual control of manual aiming movements. *Journal of Human Movement Studies*, 14, 279-291.
- Estes, W. K. (1970). Individual differences and developmental processes: The concept of learning ability. In *Learning theory and mental development* (pp. 30-32). New York: Academic Press.
- Farrell, J. N., & McDaniel, M. A. (2001). The stability of validity coefficients over time: Ackerman's (1998) model and the general aptitude test battery. *Journal of Applied Psychology*, 86(1), 60-79.
- Fisk, A. D., & Kirlik, A. (1996). Practical relevance and age-related research: Can theory advance without application? In W. A. Rogers, A. D. Fisk and N. Walker (Eds.), *Aging and skilled performance: Advances in theory and applications* (pp. 1-15). Mahwah, NJ:

Lawrence Erlbaum Associates.

- Fitts, P. M. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243-285). London: Academic Press.
- Fowler, C. A., & Turvey, M. T. (1978). Skill acquisition: An event approach with special reference to searching for the optimum of a function of several variables. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 2-38). New York: Academic Press.
- Fox, B. A. (1991). Cognitive and interactional aspects of correction in tutoring. In P. Goodyear (Ed.), *Intelligent tutoring system: Evolutions in design, teaching, knowledge, and intelligent tutoring* (pp. 149-172). Norwood, NJ: Albex Publishing Corporation.
- Gagne, R. M. (1970). *The conditions of learning* (2nd ed.). New York: Holt, Rinehart, and Winston, Inc.
- Gardner, H. (2003). Multiple intelligences after twenty years, *American Educational Research Association*.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gilhooly, K. J., & Green, A. J. K. (1989). Learning problem-solving skills. In A. M. Colley and J. R. Beech (Eds.), *Acquisition and performance of cognitive skills* (pp. 85-111). Chichester: John Wiley & Sons.
- Gillespie, M. (2003). Summary versus every-trial knowledge of results for individuals with intellectual disabilities. *Adapted Physical Activity Quarterly*, 20, 46-56.
- Goodman, J. S., Wood, R. E., & Hendrickx, M. (2004). Feedback specificity, exploration, and learning. *Journal of Applied Psychology*, 89(2), 248-262.
- Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9(6), 495-522.
- Guadagnoli, M. A., & Kohl, R. M. (2001). Knowledge of results for motor learning: Relationship between error estimation and knowledge of results frequency. *Journal of Motor Behavior*, 33(2), 217-224.
- Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing

- the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36(2), 212-224.
- Hancock, P. A., & Ganey, H. C. N. (2003). From the inverted-u to the extended-u: The evolution of a law of psychology. *Journal of Human Performance in Extreme Environments*, 7(1), 5-14.
- Hasher, L., Stoltzfuls, E., Zacks, R. T., & Rypma, B. (1991). Age and inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17(1), 163-169.
- Hodes, C. L. (1984). Relative effectiveness of corrective feedback and non-corrective feedback in computer assisted instruction in learning and achievement. *Journal of Educational Technology Systems*, 13, 249-254.
- Hogan, J. C., & Yanowitz, B. A. (1978). The role of verbal estimates of movement error in ballistic skill acquisition. *Journal of Motor Behavior*, 10, 133-138.
- Hoska, D. M. (1993). Motivating learners through CBI feedback: Developing a positive learner perspective. In J. V. Dempsey and G. C. Salas (Eds.), *Interactive instruction and feedback* (pp. 105-132). Englewood Cliffs, CO: Educational Technology.
- Johnson, T. R. (1997). Control in ACT-R and Soar. In M. Shafto and P. Langley (Eds.), *Proceedings of the 19th Annual Conference of the Cognitive Science Society*. (pp. 343-348). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Just, M. A., & Varma, S. (2002). A hybrid architecture for working memory: Reply to Macdonald and Christiansen. *Psychological Review*, 109(1), 55-65.
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*, 74(4), 657-690.
- Katz, S., O'Donnell, G., & Kay, H. (2000). An approach to analyzing the role and structure of reflective dialogue. *International Journal of Artificial Intelligence in Education*, 11(3), 320-343.
- Kieras, D. E., & Meyer, D. E. (1997). An overview of the epic architecture for cognition and performance with application to human-computer interaction. *Human-Computer*

- Interaction*, 12, 391-438.
- Kozma, R. B. (1991). The impact of computer-based tools and embedded prompts on writing processes and products of novice and advanced college writers. *Cognition and Instruction*, 8(1), 1-27.
- Kramer, A. F., Larish, J. J., & Strayer, D. L. (1995). Training for attentional control in dual-task settings: A comparison of young and old adults. *Journal of Experimental Psychology: Applied*, 1, 50-76.
- Kulhavy, R. W., & Anderson, R. C. (1972). Delay-retention effect with multiple-choice tests. *Journal of Educational Psychology*, 63(5), 505-512.
- Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1(4), 279-307.
- Kulik, J. A., & Kulik, C. C. (1988). Timing of feedback and verbal learning. *Review of Educational Research*, 58(1), 79-97.
- Laird, J., Newell, A., & Rosenbloom, P. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, 33, 1-64.
- Lavery, J. J. (1962). Retention of simple motor skills as a function of type of knowledge of results. *Canadian Journal of Psychology*, 16(4), 300-311.
- Lavery, J. J. (1964). The effect of one-trial delay in knowledge of results on the acquisition and retention of a tossing skill. *American Journal of Psychology*, 77(3), 437-443.
- Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 730-746.
- Lewis, M. W., & Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Procedural learning from examples. *Cognitive Psychology*, 17, 26-65.
- Lhyle, K. G., & Kulhavy, R. W. (1987). Feedback processing and error correction. *Journal of Educational Psychology*, 79(3), 320-322.
- Lindenberger, U., & Baltes, P. B. (1997). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin aging study. *Psychology and Aging*, 12(3), 410-

432.

- Maddox, W. T., Ashby, F. G., & Bohil, C. J. (2003). Delayed feedback effects on rule-based and information integration category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29(4), 650-662.
- Matsui, T., Kakuyama, T., & Onglatco, M. L. (1987). Effects of goals and feedback on performance in groups. *Journal of Applied Psychology*, 72(3), 407-415.
- Matthews, G., Davies, R. D., & Lees, J. L. (1990). Arousal, extraversion, and individual differences in resource availability. *Journal of Personality and Social Psychology*, 59(1), 150-168.
- McArthur, D., Stasz, C., & Zmuidzinass, M. (1990). Tutoring techniques in algebra. *Cognition and Instruction*, 7(3), 197-233.
- McCracken, H. D., & Stelmach, G. E. (1977). A test of the schema theory of discrete motor learning. *Journal of Motor Behavior*, 9(3), 193-201.
- McDowd, J. M., & Shaw, R. J. (2000). Attention and aging: A functional perspective. In F. I. M. Craik and T. A. Salthouse (Eds.), *The handbook of aging and cognition* (2nd ed., pp. 221-292). Mahwah, NJ: Erlbaum.
- McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human-Computer Interaction*, 5, 381-413.
- Merrill, D. C., Reiser, B. J., Merrill, S. K., & Landes, S. (1995). Tutoring: Guided learning by doing. *Cognition and Instruction*, 13(3), 315-372.
- Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. *The Journal of the Learning Sciences*, 2(3), 277-305.
- Metcalfe, J. (1986). Feeling of knowing in memory and problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(2), 288-294.
- Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review*, 104(1), 3-65.

- Mory, E. H. (1994). The use of response certitude in adaptive feedback: Effects on student performance, feedback study time, and efficiency. *Journal of Educational Computing Research, 11*(3), 263-290.
- Mory, E. H. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 745-783). Mahwah, NJ: Erlbaum.
- Newell, A. (1990). *Unified theories of cognition*. Cambridge, MS: Harvard University Press.
- Newell, K. M. (1991). Motor skill acquisition. *Annual Review of Psychology, 42*, 213-237.
- Newell, K. M., & Shapiro, D. C. (1976). Variability of practice and transfer of training: Some evidence toward a schema view of motor learning. *Journal of Motor Behavior, 8*(3), 233-243.
- Park, J., Shea, C. H., & Wright, D. L. (2000). Reduced-frequency concurrent and terminal feedback: A test of the guidance hypothesis. *Journal of Motor Behavior, 32*(3), 287-296.
- Peeck, J., & Tillema, H. H. (1979). Delay of feedback and retention of correct and incorrect responses. *Journal of Experimental Education, 47*, 171-178.
- Phye, G. D. (1989). Schemata training and transfer of an intellectual skill. *Journal of Educational Psychology, 81*(3), 347-352.
- Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction, 12*(3), 235-276.
- Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and performance*. Thousand Oaks, CA: Sage.
- Proteau, L., & Isabelle, G. (2002). On the role of visual afferent information for the control of aiming movements toward targets of different sizes. *Journal of Motor Behavior, 34*(4), 367-384.
- Proteau, L., Marteniuk, R. G., & Levesque, L. (1992). A sensorimotor basis for motor learning: Evidence indicating specificity of practice. *The Quarterly Journal of Experimental Psychology, 44*(3), 557-575.
- Richards, D. R. (1989). *A comparison of three computer-generated feedback strategies*. Paper

presented at the Annual Meeting of the Association for Educational Communications and Technology, Dallas, TX.

- Rogers, C. A. (1974). Feedback precision and post-feedback interval duration. *Journal of Experimental Psychology*, 102(4), 604-608.
- Rogers, W. A., & Fisk, A. D. (1991). Age-related differences in the maintenance and modification of automatic processes: Arithmetic Stroop interference. *Human Factors*, 33(1), 45-56.
- Rogers, W. A., Fisk, A. D., & Hertzog, C. (1994). Do ability-performance relationships differentiate age and practice effects in visual search? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20(3), 710-738.
- Roper, W. J. (1977). Feedback in computer-assisted instruction. *Programmed Learning and Educational Technology*, 14, 43-49.
- Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2000). Acquisition of intellectual and perceptual-motor skills. *Annual Review of Psychology*, 52, 453-470.
- Rumelhart, D. E. (1989). The architecture of mind: A connectionist approach. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 33-160). Cambridge, MA: MIT Press.
- Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin*, 95(3), 355-386.
- Sassenrath, J. M., & Yonge, G. D. (1968). Delayed information feedback, feedback cues, retention set, and delayed retention. *Journal of Educational Psychology*, 59, 69-73.
- Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4), 225-260.
- Schmidt, R. A. (June 20, 1983). This week's citation classic. *Current Contents*.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3(4), 207-217.
- Schmidt, R. A., & Wulf, G. (1997). Continuous concurrent feedback degrades skill learning: Implications for training and simulation. *Human Factors*, 39(4), 509-525.



- Schneider, V. I., Healy, A. F., & Bourne, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, 46(2), 419-440.
- Schooler, L. J., & Anderson, J. R. (1990). *The disruptive potential of immediate feedback*. Paper presented at the Proceedings of the Twelfth Annual Conference of the Cognitive Science Society.
- Schroth, M. L. (1992). The effects of delay of feedback on a delayed concept formation transfer task. *Contemporary Educational Psychology*, 17(1), 78-82.
- Schroth, M. L. (1997). Effects of frequency of feedback on transfer in concept identification. *American Journal of Psychology*, 110(1), 71-79.
- Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 5, 179-187.
- Shea, J. B., & Upton, G. (1976). The effects on skill acquisition of an interpolated motor short-term memory task during the KR-delay interval. *Journal of Motor Behavior*, 8(4), 277-281.
- Sherwood, D. E., & Lee, T. D. (2003). Schema theory: Critical review and implications for the role of cognition in a new theory of motor learning. *Research Quarterly for Exercise and Sport*, 74(4), 376-382.
- Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(4), 907-912.
- Simon, D. A., & Bjork, R. A. (2002). Models of performance in learning multi-segment movement tasks: Consequences for acquisition, retention, and judgments of learning. *Journal of Experimental Psychology: Applied*, 8(4), 222-232.
- Singley, M. K., & Anderson, J. R. (1988). A keystroke analysis of learning and transfer of text editing. *Human-Computer Interaction*, 3, 223-274.
- Smith, P. L., & Ragan, T. J. (1993). Designing instructional feedback for different learning outcomes. In J. V. Dempsey and G. C. Sales (Eds.), *Interactive instruction and feedback*

- (pp. 75-103). Englewood Cliffs, NJ: Educational Technology Publications.
- Snow, R. E. (1997). Individual differences. In R. D. Tennyson, F. Schott, N. Seel and S. Dijkstra (Eds.), *Instructional design: International perspectives* (Vol. 1, pp. 215-241). Mahwah, NJ: Lawrence Erlbaum Associates.
- Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology*, 43, 583-626.
- Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica*, 30, 276-315.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285.
- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 185-233.
- Swindell, L. K., & Walls, W. F. (1993). Response confidence and the delay retention effect. *Contemporary Educational Psychology*, 18(3), 363-375.
- Swinnen, S. P. (1990). Interpolated activities during the knowledge-of-results delay and post-knowledge-of-results interval: Effects on performance and learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 692-705.
- Swinnen, S. P., Lee, T. D., Vershueren, S., Serrien, D. J., & Bogaerds, H. (1997). Inter-limb coordination: Learning and transfer under different feedback conditions. *Human Movement Science*, 16(6), 749-785.
- Swinnen, S. P., Schmidt, R. A., Nicholson, D. E., & Shapiro, D. C. (1990). Information feedback for skill acquisition: Instantaneous knowledge of results degrades learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4), 706-716.
- Tennyson, R. D., & Breuer, K. (1997). Psychological foundations for instructional design theory. In F. Schott, N. Seel and S. Dijkstra (Eds.), *Instructional design: International perspectives: Volume 1: Theory, research, and models* (pp. 113-134). Mahwah, NJ: Erlbaum.
- Tennyson, R. D., & Schott, G. (1997). Instructional design theory, research, and models. In R.

- D. Tennyson, F. Schott, N. Seel and S. Dijkstra (Eds.), *Instructional design: International perspectives* (Vol. 1, pp. 1-16). Mahwah, NJ: Lawrence Erlbaum Associates.
- Thorndike, E. L. (1911). Laws and hypotheses for behavior. In *Animal intelligence: Experimental studies* (pp. 241-280).
- Thorndike, E. L. (1931a). The evolution of learning in general. In R. M. Elliott (Ed.), *Human learning* (pp. 162-182). New York: The Century Co.
- Thorndike, E. L. (1931b). Purposiveness and learning: Gestalt theory and learning. In *Human learning* (pp. 119-131). New York: The Century Co.
- Tuovinen, J. E., & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91(2), 334-341.
- van Hedel, H. J. A., & Dietz, V. (2004). The influence of age on learning a locomotor task. *Clinical Neurophysiology*, 115, 2134-2143.
- van Merriënboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the load off a learner's mind: Instructional design for complex learning. *Educational Psychologist*, 38(1), 5-13.
- VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology*, 47, 513-539.
- Wallace, S. A. (1996). Dynamic pattern perspective of rhythmic movement: An introduction. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 155-194). Champaign, IL: Human Kinetics.
- Watters, P. A., Martin, F., & Schreter, Z. (1997). Caffeine and cognitive performance: The nonlinear Yerkes-Dodson law. *Human Psychopharmacology: Clinical and Experimental*, 12(3), 249-257.
- Wigboldus, D. H. J., Sherman, J. W., Franzese, H. L., & van Knippenberg, A. (2004). Capacity and comprehension: Spontaneous stereotyping under cognitive load. *Social Cognition*, 22(3), 292-309.
- Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory*

- and Cognition*, 16(4), 677-691.
- Wishart, L. R., Lee, T. D., Cunningham, S. J., & Murdoch, J. E. (2002). Age-related differences and the role of augmented visual feedback in learning a bimanual coordination pattern. *Acta Psychologica*, 110, 247-263.
- Wrisberg, C. A., & Wulf, G. (1997). Diminishing the effects of reduced frequency of knowledge of results on generalized motor program learning. *Journal of Motor Behavior*, 29(1), 17-26.
- Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin and Review*, 9(2), 185-211.
- Wulf, G., Shea, C. H., & Matschiner, S. (1998). Frequent feedback enhances complex motor skill learning. *Journal of Motor Behavior*, 30(2), 180-192.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative and Neurological Psychology*, 18, 459-482.
- Zelaznik, H. N. (1996). Preface. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. ix-x). Champaign, IL: Human Kinetics.
- Zelaznik, H. N., & Spring, J. (1976). Feedback in response recognition and production. *Journal of Motor Behavior*, 8(4), 309-312.

### Author Note

This research was supported in part by Deere & Company. We thank Jerry Duncan and Bruce Newendorp for their support and advice.